Performance Discrepancy between Team Members and Motivation Gain

BY

VICTORIA M. HARMON
B.A., Northern Illinois University, 1993
M.A., Roosevelt University, 1999

THESIS

Submitted as partial fulfillment of the requirements for the degree of Doctor of Philosophy in Psychology in the Graduate College of the University of Illinois at Chicago, 2008

Chicago, Illinois

This thesis is dedicated to my husband, Steve Harmon, and my parents, Ronald and Elizabeth Smith, whose love and support made it possible.

ACKNOWLEDGEMENTS

I would like to thank my graduate advisor, Dr. James R. Larson, Jr. for his comments and contributions in the development and completion of this thesis, as well as his unwavering guidance during my graduate career at the University of Illinois at Chicago. I would also like to thank my thesis committee - Dr. Linda Skitka, Dr. Bette Bottoms, Dr. Daniel Cervone, and Dr. Robert Liden - for their thoughtful comments and suggestions that undoubtedly improved the quality of this project.

Timely completion of the data collection portion of this project could not have been accomplished without my able research assistants - Holly Heim, Jennifer Shin, and Mor Levy. Their reliability and dedication were invaluable. Finally, I would like to thank my family - especially Steven Harmon, Ronald and Elizabeth Smith, Elaine Schmucker, and Lola Harmon - whose help and support at home allowed me to complete this work.

VMH

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

TABLE OF CONTENTS (continued)

LIST OF TABLES

LIST OF FIGURES

Performance Discrepancy between Team Members and Motivation Gain

"Some of us will do our jobs well and some will not, but we will be judged by only one thing - the result" (www.vincelombardi.com, July 30, 2008).

"All I ask is that you bust your heiny on that field" (www.caseystengel.com, July 30, 2008).

Vince Lombardi was "one of football's most accomplished and respected coaches" (www.vincelombardi.com, July 30, 2008), and Casey Stengel was a baseball player, manager, or coach on 18 different teams (www.caseystengel.com, July 30, 2008). The above quotes from these two men suggest that they recognized a need to motivate their players to work hard on behalf of the team. Quotes intended to motivate team members to be "team players" are quite prevalent across all types of organizations, in sports and business alike. Leaders seem to recognize that merely being a member of team is not, in of itself, motivating. Indeed, a robust empirical finding is that team members often do not work hard on behalf of the team. Empirical demonstrations of motivation loss, a loss in motivation from working as a team member relative to working alone, are far greater in number than empirical demonstrations of motivation gain, a gain in motivation from working as a team member relative to working alone. People marvel at leaders such as Lombardi and Stengel because they motivated their team members to work hard in their teams. Perhaps this is because motivation gains are a bit of a mystery, as researchers have yet to fully explain when and why members experience that gain in motivation from working as a team member relative to working alone. Toward unraveling the mystery, the two studies reported here investigated the psychological processes behind motivation gains. An understanding of these processes would lead to successful identification of conditions under which team members work *harder* in their teams, just as the plethora of work on motivation loss

has successfully identified conditions under which team members do *not* work hard in their teams (e.g., Karau & Williams, 1993). Ultimately, understanding when and why team members experience motivation gains could provide team leaders the means to motivate their members to work hard in their teams.

The studies that have found motivation gain in the past have found that either (1) the least capable member of the team demonstrates motivation gain when the team's performance relies heavily on his or her performance; or (2) the most capable member of the team demonstrates motivation gain when the team's performance relies heavily on his or her performance. Although these motivation gains may appear on the surface to parallel one another, the extent of their similarity has yet to be empirically evaluated. The two studies reported here attempted to compare these effects directly. Specifically, their purpose was to examine whether a motivation gain by the least capable member in a conjunctive task (a team task that relies mostly on the performance of the least capable member) is fundamentally different from a motivation gain by the most capable member in an additive task (a team task that relies mostly on the performance of the most capable member), or whether, as Williams, Harkins, and Karau suggested, they are "two sides of the same motivational coin" (2003, p. 340). These studies thus tie together the currently disparate lines of research that have focused on motivation gains by either the weaker member, on the one hand, or by the stronger member, on the other, thus providing an essential step toward explaining when and why some members work harder in their teams relative to working alone.

The Kohler Effect: Motivation Gains by the Least Capable Member on Conjunctive Tasks

Gains by the least capable member were first demonstrated by Otto Kohler in the 1920's (Witte, 1989). In Kohler's experiments, members of a rowing club worked in dyads in a weight lifting task over several trials. Members were to lift the weight as many times as possible. The group's score was defined by the length of the trial, and the group trial ended when the first dyad member stopped lifting the weight (because the weight was too heavy for the remaining member to lift). Thus, the group's trial performance was dependent on the first member to stop lifting the weight, the weakest member of the dyad. The structure of Kohler's task, that the dyad's outcome was dependent on the weakest member, was later termed *conjunctive* (Steiner, 1972). Kohler found that, relative to their performance on a weight lifting task when they were alone, the least capable members worked harder in the dyad weight lifting task. A motivation gain exhibited by the least capable member in a conjunctive task is thus known as the *Kohler effect* (Witte, 1989), and has been replicated in several studies (e.g., Hertel, Kerr, & Messe, 2000; Hertel, Deter, Konradt, 2003; Hertel, Niemeyer, & Clauss, 2008; Kerr, Messe, Park, & Sambolec, 2005; Kerr, Messe, Seok, Sambolec, Lount, & Park, 2006, 2007; Lount & Phillips, 2007; Messe, Hertel, Kerr, Lount, & Park, 2002; Sambolec, Kerr, & Messe, 2007).

The Social Compensation Effect: Motivation Gains by the Most Capable Member on Additive Tasks

Gains by the most capable member have been demonstrated in work stemming from social loafing theories. Social loafing is a motivation loss, when team members work less hard in the team than they do alone. In their review of this literature, Karau and Williams (1993) suggested that perhaps social loafing cannot only be eliminated, but actually reversed under some conditions (see also Shepperd, 1993, 1995; Shepperd & Taylor, 1999). In a series of studies (Hardt, Bridgett, & Karau, 2001; Karau & Williams, 1997; Williams & Karau, 1991),

dyads worked on a brainstorming task (generating uses for a knife) either collectively (i.e., they were scored as a team) or coactively (i.e., they worked next to one another, but were scored as individuals). In the collective conditions of these studies, the members' performances were averaged or summed to create a team score, a type of group task Steiner (1972) termed *additive*. In both the collective and coactive conditions, participants were led to believe that the other dyad member would likely not perform well at the task. The researchers found that participants in the collective conditions worked harder than those in the coactive conditions, and suggested that this was because in the collective conditions the most capable member could compensate for the partner's expected lower performance by working harder. That is, because the group task was additive, the least capable member's lack of contribution to the team could be offset by an increased contribution from the most capable member. Thus, the team's performance was heavily dependent on the performance of the strongest member of the pair. Motivation gain by the most capable member in an additive task is thus known as the *social compensation effect* (e.g., Williams & Karau, 1991). A few field studies have also found evidence for the social compensation effect (e.g., Baranski, Thompson, Lichacz, McCann, Gil, Pasto, & Pigeau, 2007; Liden, Wayne, Jaworski, & Bennett, 2004), when the team task was inter-dependent. In a sleep-deprivation study, Baranski et al. (2007) found fewer fatigue effects on a computer task when participants were working with others as a team (via networked computers) than when working alone. Liden et al. (2004) found manager ratings of employee social-loafing were lower when employees expected their co-workers to loaf.

Performance Discrepancy Effects

With a few notable exceptions, studies of motivation gain have either (1) not varied the *amount* of performance discrepancy between the lower and higher performing members across

dyads, or (2) not communicated the discrepancy to the dyad members. Social compensation studies typically do not involve explicit performance feedback to the members. In the experimental studies, confederates proclaim that they do not plan to do well at the task or do not expect to do well. Thus, the amount of performance discrepancy between members is only implied by the confederate's statement and is ambiguous, at best. Performance feedback generally is included in Kohler effect studies, however few studies have provided feedback *and* varied the amount of performance discrepancy between members.

The most notable exceptions are the original studies by Kohler. As Messe et al. (2002) suggested, the members in the rowing club in Kohler's studies likely knew each other's capabilities by virtue of being in the same rowing club (they trained and raced together). Thus, in Kohler's studies, ability discrepancy varied naturally and the members knew of the discrepancy. Kohler found the greatest motivation gain in dyads with members of moderate discrepant abilities (about 1:.70 ratio; Messe et al., 2002). That is, he found an inverted-U relationship between performance discrepancy at the individual level and performance gain in the group setting. This finding was later termed the *Kohler discrepancy effect*, and was differentiated from the general Kohler effect by noting that the greatest gain came from the *moderately* less able member, compared to the gains from slightly less able or extremely less able members (Messe et al., 2002). The only studies to successfully replicate Kohler's discrepancy effect, to this author's knowledge, are two studies conducted by Messe et al. (2002). In both studies, dyad members each held a weight at arm's length above a trip rod for as long as possible, the trial length defined the team score, and the trial ended when the first member dropped the weight (thus it was a conjunctive task). In Study 1, members were informed of the performance discrepancy between them on a prior individual trial, and in Study 2, performance

discrepancy was experimentally manipulated and communicated to members via false feedback. In both of these studies, Messe and colleagues found a greater gain in performance when the discrepancy was moderate, compared to when the discrepancy was low or high.

The explanation for the Kohler discrepancy effect is largely unexplored. Although a fair amount of evidence suggests perceived indispensability and social comparison are the explanatory processes of the basic Kohler effect (Kerr et al., 2007; Hertel et al., 2008; Weber & Hertel, 2007), few attempts have been made to extend these explanations to the Kohler discrepancy effect. That is, there are no investigations into why a moderate discrepancy is more motivating than a low or high discrepancy for the least capable member in a conjunctive task. Further, if the processes operating for the least capable member in a conjunctive task also apply to the most capable member in an additive task, then a discrepancy effect may also be expected for the most capable member in an additive task, *a social compensation discrepancy effect.*

Figure 1 illustrates a discrepancy model that explains how performance discrepancies between members can (a) influence their perceived indispensability to the team and (b) influence their subsequent performance expectancies. Drawing on expectancy theories (e.g., Geister, Konradt, & Hertel, 2006; Karau & Williams, 1993; Vroom, 1964), the discrepancy model suggests that a member will improve performance if he or she feels indispensable to the team *and* expects that he or she will be able to improve performance to a level suggested by the *amount* of the performance discrepancy. Specifically, the model explains how the performance discrepancy between members influences both their perceived indispensability to the team and their expectations of improvement, and thus their motivation to work hard in their teams. Importantly, the model explains how this may occur for both the least capable member on a conjunctive task and the most capable member on an additive task. The model predicts that at

moderate discrepancy, these team members should feel both indispensable to their teams and have reasonable expectations of improvement, and thus be motivated to work hard.

The following discussion and tests of the model address only the work conditions under which participants in studies of the Kohler effect and social compensation effect have worked. The discrepancy model explains the motivational processes of the least capable member in conjunctive conditions relative to coactive conditions, and the most capable member in additive conditions relative to coactive conditions. Although the model could be extended to describe the motivation of members in other conditions (e.g., the least capable member in an additive work condition), those extensions are beyond the scope of the current project.

For the sake of simplicity, the application of the discrepancy model to the least capable member in conjunctive conditions is discussed first, followed by a discussion of the application of the discrepancy model to the most capable member in additive conditions. The same processes are considered in both cases, but they are discussed separately for ease of explanation.

In motivation gain studies, the performance of participants working collectively (scored as a team, whether additively or conjunctively) is compared to the performance of participants working coactively (scored as individuals). The discrepancy model therefore includes work condition (collective verses coactive) as a moderator. For ease of discussion, the term "member" refers to the participant in either the team condition or the coactive condition, and the term "partner" refers to the "comparison other". The comparison other is either the member's (i.e. participant's) teammate, in the team conditions, or the member's (i.e., participant's) coactor in the coactive conditions. Distinctions between the teammate and the coactor are made when appropriate. Importantly, the discrepancy model assumes that members, like those in the studies

demonstrating a discrepancy effect (Messe et al., 2002), are aware of their own and their partner's performance on the task at hand.

The Least Capable Member in a Conjunctive Task

The Kohler effect has been attributed to two motivational processes: perceived indispensability and social comparison/competition (Kerr et al., 2007; Hertel et al., 2008; Weber & Hertel, 2007). In at least one study, each accounted for 50% of the effect (Kerr et al., 2007, Study 1). In both the collective and coactive conditions, the least capable member learns of his or her partner's higher performance. This upward comparison is thought to account for higher performance when members act coactively compared to when they are alone (Kerr et al., 2007; Hertel et al., 2008; Sambolec, et al., 2007; Seta, 1982; Seta, Seta, & Donaldson, 1991). The additional boost in motivation from the coacting to the collective conditions has been attributed to perceived indispensability, the member's belief that the team's outcome is dependent on his or her level of performance (Kerr et al., 2007; Hertel et al., 2008; Weber & Hertel, 2007).

Perceived Indispensability

In a conjunctive task, the team's outcome is determined by the least capable member, thus the least capable member should feel indispensable to the team. Indispensability is closely related to instrumentality, a variable found in other models of team member motivation. Karau and William's (1993) Collective Effort Model includes instrumentality, " the degree to which high-quality performance is perceived as instrumental in obtaining an outcome" (p. 685), and Geister et al. (2006) described a model of motivation in virtual teams (the VIST model) that includes instrumentality, "the perceived indispensability of individual contributions" (p. 461). Regarding the Kohler effect specifically, Kerr et al. (2005) distinguished *possible* indispensability from *probable* indispensability, suggesting that, the participant "needs some

information indicating not just that it is possible, but that it is probable that s/he is indeed the least capable member" (p. 377). The amount of performance discrepancy between members may provide that information.

When performance discrepancy is low, the team member may view the performance differences as negligible and not recognize himself or herself as the least capable member. That is, the performance discrepancy may not be large enough to suggest to the member that it is *probable* that he or she is the least capable member. As performance discrepancy increases, that probability increases, and it should become clearer to the least capable member that he or she is indeed the least capable. Thus, in a conjunctive task, the least capable member's perceived indispensability should increase as the performance discrepancy increases.

Perceived Relative Ability

The relationship between performance discrepancy and perceived indispensability should be mediated by perceived relative ability, the member's estimate of the *ability* discrepancy between him or her and the partner. A member may attribute performance discrepancy to ability, to effort, or to both ability and effort. Members may more easily attribute small performance discrepancies to effort - the partner tried hard and/or the member did not try hard. However, as the performance discrepancy widens, it becomes more likely that the performance differences are due to differences in underlying abilities, rather than differences in effort. Members should be more certain that their respective performances are indicative of ability rather than effort as the performance discrepancy between them becomes larger. The larger the performance discrepancy, the more clear it is that one of the members is less able (for the given task) than the other, and likewise, that the other member is more able. In the discrepancy model, the member's perception of his or her ability, relative to the other member's ability, mediates the relationship

between performance discrepancy and perceived indispensability. If the member attributes the performance difference to differences in effort, as expected when the discrepancy is low, then he or she is less likely to conclude that he or she is the least capable member. That is, the probability of the member viewing himself or herself as the least capable, and therefore indispensable in the conjunctive task, is lower when the difference is attributed to effort rather than ability.

The relationship between perceived relative ability and perceived indispensability is moderated by work condition such that there is a negative relationship when the members work conjunctively but not when they work coactively. Under coactive work conditions, the members are rewarded as individuals, not as a team. In the absence of a team outcome, each member's individual outcome is completely dependent on his or her own performance and not at all dependent on the performance of others. Perceived relative ability should thus have no effect on perceived indispensability in coactive conditions.

Performance Improvement Expectancy

In accordance with other expectancy-based models (e.g., Geister, et al., 2006; Karau & Williams, 1993; Vroom, 1964) the discrepancy model suggests that motivation is influenced by the member's performance improvement expectancy, the member's belief that he or she can improve his or her performance to a specific level. Performance improvement expectancy is closely related to self-efficacy and expectancy (Earley & Lituchy, 1991), which have been included in other models of team member motivation (Geister, et al., 2006; Karau & Williams, 1993). The Collective Effort Model includes expectancy, "the degree to which high levels of effort are expected to lead to high levels of performance" (Karau & Williams, 1993, p. 685), and the VIST model includes self-efficacy, "the perceived capability to fulfill the tasks required in a

team" (Geister et al., 2006, p. 461). Bandura and Cervone (1983) found that providing both goals and feedback to participants resulted in higher performance compared to when no goals and/or no feedback was provided. Importantly, in the goals/feedback condition, self-efficacy was predictive of performance.

Performance discrepancy between members not only provides feedback to the member about his or her performance relative to others, but can also suggest a level of performance improvement for the least capable member (Seta, 1982; Seta, et al., 1991). There are a number of ways in which performance discrepancy may suggest performance improvement to the least capable member (Kerr et al., 2007). First, the upward social comparison the least capable member makes when comparing his or her performance to the partner may induce competition. Several studies suggest that the least capable members are motivated by competition, trying to meet or beat their partner's performance level (Kerr et al., 2007; Lount & Phillips, 2007; Sambolec et al., 2007). Second, this upward social comparison may suggest an adjustment of personal goals upward due to a "motive to be no worse" than others (Stroebe, Diehl, & Abakoumin, 1996). That is, the least capable member adjusts his or her personal performance goals to meet the most capable member's level, to avoid being the "worst" member. A third reason involves the motive to obtain or maintain favorable evaluations from peers (Seta, 1982). Again, the performance level of the most capable member suggests the level of improvement needed from the least capable member. Finally, in a team setting, the least capable member may want to avoid hurting his or her team, and thus may strive to at least match the other's performance. This may be especially the case when the team's performance depends on that of the least capable member, such as when the task is conjunctive. Thus, there are a number of

ways in which the performance discrepancy between the most capable and least capable members may serve as a basis for a level of improvement for the least capable member.

When performance discrepancy is large, the suggested level of improvement is high, and the amount of performance improvement needed to reach that level is large. These larger improvements should be perceived as less achievable than the smaller improvements that would be implied by smaller performance discrepancies. Thus, for the least capable member, performance improvement expectancy should decrease as performance discrepancy increases. That is, as discrepancy increases, expectation of being able to meet the suggested level of improvement decreases.

Although this may hold for the least capable member paired with either a coactor or with a teammate, the least capable member paired with a teammate may have an additional motive over that of one paired with a coactor. Specifically, in a team setting with a conjunctive task, the least capable member may be motivated to reach the partner's level of performance to avoid hurting the overall team performance. If the least capable member brings his or her performance up to the level of the partner, then he or she can not be accused of hurting the team's performance. This additional motive suggests the performance discrepancy is more likely to imply an improvement level for the least capable member in a conjunctive task than for the least capable member in a coactive task.

Performance

According to the model, the level of performance discrepancy that leads to (1) the member's perception that he or she is indispensable to the team and (2) the member's expectation that he or she can improve performance to the partner's level, should in turn lead to (3) an increase in performance from working alone to working as a team member. When performance

discrepancy between team members is low, the least capable member may not see himself or herself as truly the least capable member, and thus may not perceive himself or herself as particularly indispensable to the team, so there should be low motivation to improve. When the performance discrepancy between team members is high, the least capable member should have low expectations of improving his or her performance to the partner's level, so there should be low motivation to improve. By contrast, the motivation of the least capable member in a conjunctive task should be relatively high when the discrepancy between members is moderate. At a moderate discrepancy, the least capable member in a conjunctive task (1) is more likely attribute performance differences to ability and thus see himself or herself as indispensable to the team and (2) expect that he or she can improve performance to the partner's level. These processes should lead to the greatest increase in performance from working alone to working as a team member for the moderately least capable member in a conjunctive task, a prediction commensurate with the Kohler discrepancy effect (Messe et al., 2002). The Kohler discrepancy effect specifies that the Kohler effect, motivation gain of the least capable member in a conjunctive task, is greater when the performance discrepancy between members is moderate, than when the discrepancy is low or high.

Performance of the least capable member in coactive conditions should be solely a function of performance improvement expectancy, because performance discrepancy should not influence perceived indispensability to the team when there is no team. However, even under coactive conditions, the most capable member's performance level may suggest an improvement level to the least capable member (via upward social comparison). At low levels of discrepancy, the least capable member should expect to easily improve performance to the partner's level, and at high levels of discrepancy, improving performance to the partner's level should be viewed as

very difficult. In contrast, at moderate levels of performance discrepancy, the partner's performance level suggests a reasonable amount of improvement, and thus should be most motivating. The least capable member in *either* a conjunctive task *or* a coactive task should thus be more likely to improve performance under moderate discrepancy conditions than under low or high discrepancy conditions. However, the performance improvement of the least capable member under moderate discrepancy conditions should be *greater* in a conjunctive task than in a coactive task because the member in a conjunctive task has the additional motive to avoid hurting the team's outcome with low performance.

Personal Goals

Finally, the discrepancy model suggests that performance is preceded by the member's personally adopted performance goal. Studies examining self-efficacy or performance expectancy typically find that the relationship between efficacy/expectancy and performance is mediated by personal performance goals (e.g., Earley & Lituchy, 1991; Tolli & Schmidt, 2008; Vrugt & Koenis, 2002; Vrugt, Oort, & Zeeburg, 2002). Because perceived indispensability is a key motivator behind motivation gain (Kerr et al., 2007; Hertel et al., 2008; Weber & Hertel, 2007), personal goals are expected to be influenced by both performance improvement expectancy and perceived indispensability, which in turn leads to performance. That is, the influence of perceived indispensability and performance improvement expectancy on performance should be mediated by the member's personally adopted performance goal.

Summary of Least Capable Member

The discrepancy model predicts a Kohler discrepancy effect like that previously demonstrated (Messe et al., 2002). The least capable member, under both conjunctive and coactive work conditions, should be most motivated when performance discrepancy is moderate.

However, the motivation under moderate discrepancy should be higher for the least capable member in a conjunctive task, compared to in a coactive task, because in a conjunctive task the least capable member should perceive himself or herself as indispensable to the team and want to avoid hurting the team with low performance.

The Most Capable Member in an Additive Task

The explanations for the Kohler effect as described in the discrepancy model seem applicable to the social compensation effect as well. In both the collective and coactive conditions, the most capable member learns of his or her partner's lower performance. This downward comparison should not influence motivation of the most capable member in the coactive conditions, but should influence motivation of the most capable member in additive conditions by implying a level of improvement needed to compensate for the least capable member. An additional motivator for the most capable members in additive conditions, compared to those in the coactive conditions, is perceived indispensability, the member's belief that the team's outcome is dependent on his or her level of performance. These processes are next explained.

Perceived Indispensability

Performance discrepancy should influence the most capable member's perceived indispensability when the task is additive. Although the contribution of every member has some impact on the team's outcome in an additive task, the most capable member's contribution has *more* of an impact simply because it is *larger* than the others' contributions. For the most capable member in an additive task, it should become more clear that he or she is "carrying" the team, as the difference in his or her contribution, relative to that of the other members, becomes larger. Thus, for the most capable member in an additive task, perceived indispensability to the

team should increase as the performance discrepancy between him or her and the other member widens.

Perceived Relative Ability

Just as the least capable member may attribute small performance differences to effort rather than ability, so too may the most capable member. For example, the most capable member may assume that the partner did not try hard when the performance discrepancy was small. As performance discrepancy widens, members should be more certain that their respective performances are indicative of ability. In the discrepancy model, the member's perception of his or her ability, relative to the other members, mediates the relationship between performance discrepancy and perceived indispensability. If the member attributes the performance difference to differences in effort, as expected when the discrepancy is low, then he or she is less likely to conclude that he or she is the most capable member, and thus less likely to perceive himself or herself as particularly indispensable to the team.

The relationship between perceived relative ability and perceived indispensability is moderated by work condition, such that there is a positive relationship when the members work additively but not when they work coactively. Under coactive work conditions, the members are rewarded as individuals, not as a team. In the absence of a team outcome, each member's individual outcome is completely dependent on his or her own performance, and not at all dependent on the performance of others. Perceived relative ability should thus have no effect on perceived indispensability in coactive conditions.

Performance Improvement Expectancy

Under additive work conditions, the performance discrepancy between members may suggest a level of improvement to the most capable member, if the most capable member wishes

to compensate for the least capable member. In the VIST model, self-efficacy is defined as the ability to "fulfill the tasks required in a team" (Geister et al., 2006, p. 461). For the most capable member in an additive task, effectively fulfilling the tasks required in a team may entail effectively compensating for the least capable member. If the least capable member is contributing much less than the most capable member (i.e., the performance discrepancy is large), then the most capable member must contribute quite a bit (i.e., improve a great deal) to compensate for the least capable member's small contribution. Higher levels of discrepancy thus imply higher levels of compensation. Higher levels of compensation entail higher levels of performance improvement on the part of the most capable member, which should be perceived as fairly difficult to attain. That is, the most capable member should have a low expectancy of improving performance to the level required to effectively compensate for the partner, when the performance discrepancy between them is high. At lower levels of performance discrepancy, the difference between the least capable and most capable members' contributions is not as great, and the amount of compensation required of the most capable member is lower. Lower levels of compensation entail lower levels of performance improvement on the part of the most capable member, which should be perceived as fairly easy to attain. That is, the most capable member should have a high expectancy of improving performance to the level required to effectively compensate for the partner, when the performance discrepancy between them is low. Thus, because the most capable member must make larger performance improvements to compensate when the performance discrepancy is large, and larger performance improvements are more difficult to achieve, the most capable member's performance improvement expectancy should decrease as performance discrepancy increases.

The most capable member in a coactive task should not be concerned with compensating for the partner's lower performance because there is no team outcome. The most capable member in a coactive task therefore is left with only a downward social comparison when learning of the partner's lower performance. This downward comparison may imply a small performance improvement to the most capable member if the performance discrepancy is low. Stroebe et al. (1996) suggested that both the most capable member and least capable member have a motive "to be better than others" (p. 57). The most capable member may increase performance to keep his or her place as the "most" capable member. However, this likely would occur only when the discrepancy is small. As the discrepancy widens, the most capable member's status is more secure, and thus he or she may be less motivated to improve performance. Thus, low levels of discrepancy may imply a small performance improvement to the most capable member in a coactive task, but moderate to high levels of discrepancy are less likely to imply any performance improvement to the most capable member in a coactive task. Because the amount of discrepancy does not suggest any significant level of improvement to the most capable member in coactive tasks, the amount of discrepancy should not influence the most capable member's expectancy of performance improvement. Thus, there should be no differences in performance improvement expectancy across discrepancy levels for the most capable member in a coactive task.

Performance

According to the discrepancy model, the level of performance discrepancy that leads to (1) the member's perception that he or she is indispensable to the team and (2) the member's expectation that he or she can improve performance enough to compensate for the partner, should in turn lead to (3) an increase in performance from working alone to working as a team

member. When performance discrepancy between team members is low, the most capable member may not see himself or herself as much more capable, and thus may not perceive himself or herself to be especially indispensable to the team. When the performance discrepancy between team members is high, the most capable member should have low expectations of improving his or her performance enough to compensate for the least capable member. The discrepancy model thus predicts that motivation gain of the most capable member in an additive task should be greatest when the discrepancy between members is moderate. At a moderate performance discrepancy, the most capable member (1) is more likely to see himself or herself as indispensable to the team and (2) expect that he or she can improve performance enough to compensate for the least capable member. These processes should lead to the greatest increase in performance from working alone to working as a team member for the moderately most capable member in an additive task. The performance predictions of the discrepancy model for the most capable member in an additive task are thus the same as those for the least capable member in a conjunctive task. Most importantly, the discrepancy model predicts a social compensation discrepancy effect parallel to the Kohler discrepancy effect.

Under coactive conditions, the performance discrepancy between members should not suggest an improvement level to the most capable member, except perhaps at low discrepancy, in which the most capable member may want to slightly improve to keep the "most capable" status. Further, perceived indispensability to the team would not be a motivator when there is no team. Therefore, performance discrepancy should not influence the performance of the most capable member in a coactive task.

Personal Goals

Finally, just as for the least capable member, the most capable member's personally adopted performance goal should mediate the influences of perceived indispensability and performance improvement expectancy on his or her performance.

Summary of the Most Capable Member in an Additive Task

The discrepancy model predicts a social compensation discrepancy effect parallel to the Kohler effect. The most capable member in an additive task should be most motivated when performance discrepancy is moderate due to perceived indispensability and performance improvement expectancy. Under coactive work conditions, performance discrepancy should not influence performance of the most capable member.

Overview of Studies

The discrepancy model makes a number of predictions for the least capable and most capable members in conjunctive and additive conditions, respectively, in comparison to a coactive condition. Two studies were conducted to (a) test for performance discrepancy effects (i.e., Kohler discrepancy effect and social compensation discrepancy effect), and (b) test the underlying processes of motivation gain described in the discrepancy model. Next, an overview of the task used in the studies is provided, followed by an overview of the studies' design. Detailed reports of each study follow the overview. The paper concludes with a general discussion.

The Vigilance Task

The Kohler effect is often studied using a physical persistence task, for example, holding a weight above a trip rod for as long as possible (Hertel, et al., 2000; Messe, et al., 2002; Stroebe, et al., 1996). In this type of task, the motivated participant holds the weight longer than the unmotivated participant. That is, if motivated, the participant will exert more effort and will

persist at the task, even after becoming fatigued. Participants in the current studies, on the other hand, performed a visual vigilance task. In the physical persistence task, the motivated participant resists the desire (due to physical fatigue) to relax and lower the weight. In the visual vigilance task, the motivated participant resists the desire (due to cognitive fatigue) to relax and become distracted from the task. This task was used for both studies, and is described in detail under Study 1 procedures.

Other cognitive tasks have been used in the study of motivation gain. For example, social compensation studies typically use a brainstorming task (e.g., Williams & Karau, 1991). Recently, the Kohler effect was found using simple calculation math problems (Lount & Phillips, 2007) and in a computer-based task (Hertel, et al., 2003; Hertel, et al., 2008). Although these tasks certainly require cognitive effort, it is unclear how dependent they are on the participant's ability to concentrate. The vigilance task used in the studies reported here was modeled after a task designed to assess the ability to sustain attention (Manly, Robertson, Galloway, & Hawkins, 1999; Robertson, Manly, Andrade, Baddeley, & Yiend, 1997), and the task was billed to participants as a measure of concentration.

Overview of Studies' Design

Participants in both studies worked in dyads either coactively or collectively. In the coactive conditions participants received performance information about one another, but they were always scored as individuals. In the collective condition, participants received performance information about one another, and on half of the trials they were scored as a team. In all work conditions, participants performed the visual vigilance task (i.e., detecting target stimuli on a computer screen) over four trials. After both the first and second trials, in which all participants were scored as individuals, all participants learned of their own performance, the performance of

their partner, and the amount of discrepancy between them. In the third and fourth trials, participants in the coactive conditions continued to be scored as individuals, while those in the collective conditions were scored as a team. The level of performance discrepancy between members was manipulated such that participants were led to believe that there was a either a large, moderate, or small performance discrepancy between them. Participants in the collective conditions of the first study were led to believe that they were the least capable member in a conjunctively scored task. Participants in the collective conditions of the second study were led to believe that they were the most capable member in an additively scored task.

Study 1: The Least Capable Member in a Conjunctive Task

Study 1 examined the motivational processes of the least capable member in a conjunctive task as described by the discrepancy model. According to the discrepancy model, motivation should be highest for the least capable member in a conjunctive task under conditions of moderate performance discrepancy (i.e, Kohler discrepancy effect).

Description of Performance Trials

Participants performed the task in four different trials. The first trial was intended as a baseline measure of their ability. All of the participants were scored as individuals on Trial 1, and none had yet learned of their partner's performance. After Trial 1, participants learned that their score on Trial 1 was lower than their partner's scores on that same trial (i.e., the performance discrepancy manipulation). On Trial 2, all the participants continued to be scored as individuals, but under the knowledge that they had scored lower than their partner on Trial 1. After Trial 2, participants were told that their score on Trial 2 was again lower than their partner's score. On Trial 3, the participants in the conjunctive conditions were scored as a team, such that the team score was equal to the score of the first person to stop during the trial. The

computer was programmed to lead participants to believe they stopped before their partner, and showed the team score as being equal to the participant's score. This is what made the team task a conjunctive task. By Trial 4, then, the participants in the conjunctive conditions were under the impression that their partner had outperformed them on the three prior trials. Further, by the fourth trial the participants in the conjunctive conditions had experienced the conjunctive scoring by seeing that the team score was equal to their score (because they had apparently stopped the trial before their partner). Those in the coactive conditions, by contrast, were scored as individuals on Trials 3 and 4, and did not learn of their partner's performance on those trials.

Hypotheses

In their replication of Kohler's discrepancy effect, Messe at al. (2002) also had pairs of participants perform four trials, the first two scored as individuals and the second two scored as a team. These researchers found Kohler's discrepancy effect only in the fourth trial, after the participants had experienced the conjunctive scoring on the third trial. Drawing from those findings, performance differences between experimental conditions in the present study were expected only on the fourth trial, after those in the conjunctive conditions had experienced the conjunctive nature of the task. The effect of work condition (conjunctive or coactive), and the interactive effect of work condition and performance discrepancy (low, moderate, or high) on performance should be most evident in Trial 4. Further, their effects should also be evident in the change in performance from Trial 2 (after seeing the performance discrepancy for a second time) to Trial 4 (after experiencing the conjunctive task). The first two hypotheses reflect these predictions. Hypotheses three through six are predictions concerning the motivational processes suggested by the discrepancy model.

Hypothesis 1: Performance on the fourth trial will be greatest in the conjunctive condition when performance discrepancy is moderate.

Hypothesis 2: The greatest increase in performance from the second trial to the fourth trial will occur in the conjunctive condition when performance discrepancy is moderate.

Hypothesis 3: Perceived relative ability will decrease as performance discrepancy increases.

Hypothesis 4a: In the conjunctive conditions, perceived indispensability will increase as performance discrepancy increases.

Hypothesis 4b: In the coactive conditions, perceived indispensability will not vary across levels of performance discrepancy.

Hypothesis 5: Performance improvement expectancy will decrease as performance discrepancy increases.

Hypothesis 6: Personal performance goals for the fourth trial will be greatest in the conjunctive condition when performance discrepancy is moderate.

In addition to testing these hypotheses individually, the discrepancy model as a whole will also be tested via path analysis.

Method

Visual Vigilance Task

Two key features were incorporated into the development of the visual vigilance task to assure the task required participants' concentration. The first feature is the "response" of a non-response. The task involves a series of numbers (single digits 0 through 9), presented one at a time, on a computer screen. The participants compare each presented number to a given target. If the presented number does *not* match the target, the participant responds by pressing the space

bar. However, if the presented number *does* match the target, the participant does nothing, he or she "responds" by *not* pressing the space bar. This feature of the task, withholding a response when a match occurs, was modeled after a task used to assess ability to sustain attention (Manly, et al., 1999; Robertson, et al., 1997), which "requires the participants remain sufficiently attentive to their responses such that, at the appearance of the target, they can substitute the directly antagonistic response of not pressing" (Manly et al., 1999, p. 662).

The second key feature of the task involves participants keeping in memory the target to which they compare the presented number. The target is a number *previously*, not currently, displayed on the computer screen. That is, participants have to compare each stimulus to a target number that was previously displayed. For example, at the beginning of a performance trial, a "2" might be displayed next to the word "target." After a few seconds, that "2" disappears and another number, perhaps an "8," appears. Then the stimulus numbers begin to be presented in the center of the screen. Participants compare the presented number to the previously displayed target (2), not the currently displayed target (8). Thus, if any number other than a 2 is presented, the stimulus and target (2) do *not* match, and the participant presses the space bar. If a 2 is presented, the stimulus and the target (again, 2) *do* match, and the participant does *not* press any key; he or she withholds a response.

After some time, the target currently displayed on the screen changes; for example the "8" may change to a "1." The participants must then compare the subsequently presented stimulus numbers to the "8," the previously displayed target (not to the target currently displayed, 1). If any number other than an 8 is presented, it is *not* a match to the target (8) and the participant presses the space bar. If the presented number is an 8, however, it *does* match the target, and the participant withholds a response. Thus, a target number begins to function as a

target only when it disappears from the screen. This memory feature of the task increases the participant's cognitive load by requiring the participant to keep the comparison target in memory. Such a task is referred to as a successive task, one in which "subjects need to compare current input with a standard retained in working memory in order to separate signals and noise" (Huey & Wickens, 1993, p. 148). This stands in contrast to a simultaneous task, in which "all of the information needed to distinguish signals from nonsignals is present on the stimuli themselves" (Huey & Wickens, 1993, p. 148). Performance on a simultaneous task is less affected by distracting thought than performance on a successive task (Smallwood, Davies, Heim, Finnigan, Sudberry, O'Connor, & Obonsawin, 2004). To perform well on a successive task, therefore, the participant must carefully concentrate.

The computer screen is black, and the target and stimulus numbers are presented in white. The first target (e.g., 2) is shown for 2000 ms before it changes (e.g., to 8). The presentation of the stimulus numbers begin when the first target changes (e.g., when 2 changes to 8). The presentation speed of the stimuli increases steadily, such that at first the stimulus is displayed for 2100 ms, but over the first 3.35 minutes of the trial, the stimulus presentation rate decreases to a display time of 700 ms. The presentation speed then remains steady at one number every 700 ms. The inter-stimulus interval is zero. The order of the targets is the same for every trial and for every participant. The order of the stimuli is different for every trial, but the same for every participant. Thus, the trials are the same across participants. The targets remain on the screen for a varying number of stimulus presentations, ranging from one to twenty. Just before the target changes, it turns a different color (purple) as a signal that it is about to change.

After each stimulus presentation, the participant either presses the space bar or does not press a key, and the computer tells the participant whether his or her response/non-response was

correct or incorrect. The word "wrong" flashes in red when the response is incorrect and the word "right" flashes in green when the response is correct. There are two types of correct responses, or "hits," pressing the space bar when the target and stimulus are different, and not pressing any key when they are the same. There are also two types of errors. A "miss" is pressing the space bar when the target and the stimulus match, and a "false alarm" is not pressing the space bar when the target and stimulus do not match. Although the computer recorded hits, misses, and false alarms, responses were referred to as simply "right" or "wrong" to the participants.

Displayed in the upper right-hand corner of the computer screen is a tally of the number of stimuli presented and the number of "strikes" the participant has accumulated. A "strike" is an error made in the last five stimulus presentations. Once the participant accumulates three strikes, or three errors in five stimulus presentations, the trial ends. Strikes can be erased such that, if the participant makes an error and thus has a strike, but responds to five more stimulus numbers without an error, that strike is removed. This allows participants to recover from errors.

Once three strikes are accumulated, the presentation of the stimuli stops. A "Get Scores" button then appears, which when clicked displays the participant's scores for that trial. The total number of turns (stimulus presentations), the number of correct responses, and the number of incorrect responses are displayed. Once the participant has reviewed the scores, he or she exits the trial by clicking a button that appears when the scores appear.

A pilot test (described in Appendix A) was conducted to assure that increased effort on the vigilance task is reflected in increased performance. That study found that participants who were told to increase effort also had increased performance scores, compared to participants who

were not told to increase effort. Thus, participants who exert more effort on the task should out-perform participants who exert less effort.

Participants

194 participants, of whom 96 were male, were recruited from the UIC subject pool. Their average age (of those who reported age, n = 184) was 19.53 years. The study used a 2 (work condition; conjunctive, coactive) x 3 (performance discrepancy; low, moderate, high) x 2 (performance trial) factorial design, with the last factor repeated. With the exception of balancing gender within conditions, participants were randomly assigned to the various experimental conditions. Besides the six between-subject conditions, a seventh condition in which individuals performed the task over four trials, with no manipulations, was included to provide a control for practice and fatigue effects (described in more detail in the results section). An additional 37 participants, of whom 21 were male (average age 19.16 years) were recruited for this control condition.

Procedure

Participants came to the lab in same-sex pairs and sat at the same table, where the experimenter gave all of the verbal instructions. She told them that they would be asked to perform a visual vigilance task over several trials on the computer. They were told that they would perform the task in separate computer rooms, but the computers were networked. After completing the informed consent process, the experimenter told the participants in the conjunctive conditions that they would be working with the other participant as a team for some of the trials, and asked participants to complete a "Getting to Know Your Partner" form. The form required the participants to ask each other their names, 10 questions about each other (e.g.,

What is your favorite color? What is your favorite food?), and then decide together on a team name. Participants in the coactive conditions did not fill out this form.

The experimenter then told the participants that the visual vigilance task they would be performing is a good measure of how long one can concentrate. The experimenter also told the participants that companies may use vigilance tasks like this one to hire employees. For example, an airline company may use a vigilance task to hire airline pilots. Similar instructions were used by Helton, Dember, Warm, and Matthews (1999), and were intended to make the task meaningful to participants.

The experimenter then gave detailed verbal instructions on how to perform the task, followed by a demonstration of a practice trial on a computer. The practice trial consisted of 10 stimulus numbers. Presentation of the numbers in the practice trial continued regardless of the number of errors made (i.e., the presentation of stimuli did not stop after 3 strikes occurred). After the demonstration, each participant performed the practice trial himself or herself, in separate computer rooms. Once they completed the practice trial, the participants moved back to the table in the instruction room.

At the end of each trial the computer reported to the participant his or her total number of turns (number of stimuli presented, or the length of the trial), number of right responses and number of wrong responses. The "score," as referred to below, was the length of the trial (the total number of turns).

The experimenter explained to the participants that they would encounter five trials - a learning trial, two individual trials, and two team trials (for those in the coactive and control conditions, the last four trials were referred to as performance Trials 1 through 4). The experimenter explained that the learning trial would consist of a maximum of 50 numbers.

Further, if they forgot what the target was, participants could press the "Enter" key and the target would briefly appear on the screen. However, they could do this during the learning trial only. The experimenter then told them that after the first trial and after the second trial ("first/second individual trial" for teams, "first/second performance trial" for coactors), the participants would have the opportunity to view scores from past participants and each other's scores. She explained to those in the team conditions that the next two trials after that would be the team trials, and to those in the coactive conditions that the next two trials after that would be performance Trials 3 and 4.

The experimenter explained that the participants would be rewarded for good performance on the task with tickets for a lottery to be held at the end of the semester[1]. If the participants won the lottery, they received $50.00 each. Participants earned one lottery ticket for every 50 points scored. The experimenter told them that lottery tickets would be awarded based on performance on only the last two trials. Thus, for the teams, the participants believed they would receive lottery tickets on only the team trials, based on their team score on those trials. In other words, they thought tickets were awarded to the team, not to the individual participants (each team member would get 50.00 if the team won). Participants in the coactive conditions were told they would receive lottery tickets on only the last two trials based on their individual scores. In other words, they thought tickets were awarded to them as individuals, completely independent from the other participant. The participants in the coactive conditions were told that the reason for awarding lottery tickets based on only the last two trials was because people typically get rewarded for tasks they are very familiar with. So, they were receiving tickets on only the last two trials, after they had become familiar with the task. The participants filled out a

form with contact information so they could be contacted at the end of the semester with more information about the study and the results of the lottery.

The participants in the team conditions then learned how the team score would be figured. They were told that, when one partner's trial ended, the other partner would not be able to continue, and the team score would be equal to the score of the partner who stopped first. The experimenter then demonstrated how the team trials work on the computer.

Participants then went to their separate computer rooms. They were instructed to wear earmuffs to avoid distractions, and the overhead light was turned off to reduce glare on the screen (there was still light in the room from either a lamp or a window). The experimenter shut the door and the participant proceeded through the rest of the experiment by following the directions on the computer. Table 1 lists the order of operations on the computer that participants in the conjunctive conditions encountered, and Table 2 lists the order of operations for those in the coactive conditions.

When the participants completed all five trials, they then filled out a final questionnaire, also on the computer. They were then thanked by the experimenter and dismissed. Participants were sent a debriefing form at the end of the semester.

Work condition manipulation. During the team trials, the participant's ID code (a code he or she entered into the computer before the first trial) and the partner's ID code (a fake code, every participant's "partner" had the same ID), appeared in the lower left corner. When one participant stopped the trial, his or her ID code disappeared. This was modeled after Kerr et al. (2006). In the demo program shown by the experimenter, the ID codes were simply "Partner C" and "Partner D". In the first demonstration trial, the experimenter showed the participants what it would look like if the participant stopped before his or her partner. The experimenter performed

the task for 20 numbers, then stopped responding, allowing three strikes to quickly accumulate, thereby ending the trial. The experimenter noted that her ID code (Partner C) had disappeared, because she had three strikes, but that her partner's ID code (Partner D) was still visible because that person had not reached three strikes. After the trial, the participants saw the experimenter's score (Partner C), the "partner's" score (Partner D) and the team score. The experimenter noted that the "partner" had been prevented from continuing because she, the experimenter, had stopped the trial, and the team score was equal to her (the experimenter's) score because she was the first to stop. On the second demonstration trial, the experimenter showed the participants what it would look like if the partner stopped before the participant. After 15 numbers had been presented, the "partners" ID code (Partner D) disappeared. The experimenter noted that this was because the partner had three strikes, but that her ID code (Partner C) was still visible because she had not reached three strikes. When the partner's ID code disappeared, the experimenter was cut-off from continuing the trial because the "partner" had stopped. The numbers stopped being presented and a note said the trial was over because the partner had stopped. After the trial, the participants saw the experimenter's score (Partner C), the "partner's" score (Partner D) and the team score. The experimenter noted that she had been cut-off from continuing because the partner had stopped the trial, and the team score was equal to her partner's score because the partner was the first to stop. During the actual trials, participants were always led to believe that they were the first to stop during the team trial, as the partner's ID code never disappeared (the partner never appeared to reach three strikes). The participant's score was always genuine, and the partner's ostensible score was one plus the participant's score (to give the impression that the partner was cut-off from continuing the trial). The team score was always equal to the participant's score.

For those in the coactive and control conditions, the participants' ID codes were not displayed during any of the trials, and, in the demo program, the experimenter demonstrated how the lottery tickets worked. In the first trial of the demonstration, the experimenter performed the task for 25 numbers, then stopped responding, allowing three strikes to quickly accumulate and end the trial. She showed the participants her score and noted she had no lottery tickets because she did not reach 50 points. In the second trial of the demonstration, the "score" started at 50, and the experimenter performed the task for 15 numbers, then allowing three strikes and ending the trial (for a "score" of 68). Again, she showed the participants her score and noted she had only one lottery ticket because she had 50 points, but did not have two lottery tickets because she had not reached 100 points. She noted that no fractional lottery tickets would be awarded.

Discrepancy manipulation. After the first individual trial for those in the conjunctive conditions, (called the "first performance trial" for those in the coactive conditions), participants were shown a list of scores, ostensibly from past participants and from their "partner" (or, the "person in the next room" for those in the coactive conditions). The list contained 34 scores, from highest (1st place) to lowest (34th place), and the participant's score was always in the 17th place. The other scores in the list, including the partner's ostensible score, were all computed as a fixed percentage of the participant's score (see Figure 2). The participant was always led to believe that his or her score was lower than his or her partner's score. In the low discrepancy conditions, the partner's score was 1.05 times the participant's score, and the partner's place in the list was 15 (for example, if the participant's score was 200, the partner's score was 210). In the moderate discrepancy conditions, the partner's score was 1.17 times the participant's score, and the partner's place in the list was 11 (for example, if the participant's score was 200, the partner's score was 234). In the high discrepancy conditions, the partner's score was 1.84 times the

participant's score, and the partner's place in the list was 2 (for example, if the participant's score was 200, the partner's score was 368).

After the second trial ("second individual trial" for those in the conjunctive conditions, "second performance trial" for those in the coactive conditions), participants were shown a list of scores constructed in exactly the same way as after the first trial. Except for the score values in the list, which were now based on the participant's second trial score, the list was exactly the same as the list shown after the first trial (e.g., same "participants" in the same places on the list).

After the first trial and after the second trial, the participant clicked a button to view the list of scores. When they clicked that button, a "wait" screen appeared which instructed them that their partner was not yet finished, and they had to wait to view the scores. The wait time was commensurate with the discrepancy level. Specifically, the wait time was the discrepancy amount times 7 ms. For example, if the participant was in the low discrepancy condition and had a score of 200, the discrepancy was 10 (partner's score was 210), and the wait time was 70 ms (10 times 7ms).

A pilot study (described in appendix B) was conducted to ensure that participants viewed the three levels of discrepancy as low, moderate, and high, respectively. The results of the pilot study showed that, indeed, participants in the low discrepancy condition viewed the performance difference between them and their partner as low, participants in the moderate discrepancy condition viewed that performance difference as moderate, and participants in the high discrepancy condition viewed that performance difference as high.

Measures. The main performance dependent variable was the score on Trial 4, which the computer recorded. After receiving the feedback about ostensible others' scores, the participants completed the performance improvement expectancy measure on the computer. They received

the feedback two times (once after the first trial and once after the second trial), and thus completed the performance improvement expectancy measure two times.

After the fourth and final trial, all participants responded to a questionnaire presented on the computer. The questionnaire measured perceived indispensability, perceived relative ability, and personal goals. Also included were manipulation checks, and measures of control variables such as perceived value of the lottery and video game experience. Appendix C lists all of the questionnaire items.

Control condition. Individuals in the control condition performed the task for four trials, and, like those in the coactive conditions, were scored as individuals on all four trials. However, unlike the coactive conditions, participants came to the lab one at a time (rather than with a coactor), and they did not get any information about others' performance on the task. Otherwise, they received the same instruction as those in the coactive conditions. Just as those in the experimental conditions, participants in the control condition were told they would be rewarded with lottery tickets on the last two trials. Table 3 lists the order of operations on the computer for participants in this control condition. These participants also answered most of the questions that participants in the other conditions answered. The exceptions are noted in Appendix C.

Results

Data Cleaning for the Trial Scores

The total number of stimuli presented during a trial was the trial score. This score is equivalent to the length of the trial, a measure of how long participants persisted at the task.

Trial scores were screened for univariate outliers and missing values. None of the cases were identified as univariate outliers on Trial 1, and fourteen cases were identified as univariate outliers on either Trial 2, 3, or 4 (z-score > 3.29, $p < .001$, two-tailed test; Tabachnick & Fidell,

2001). None of these outliers were dropped from analyses because they did not have extreme scores on the first trial. That is, their baseline performance was not extreme.

One case was dropped due to a technical error (the wrong condition number was entered in the computer). Six cases were dropped because they were missing data on one or more of the four performance trials. Of these six cases, five were in the conjunctive condition (two in the high discrepancy condition, three in the moderate condition) and one was in the coactive/high discrepancy condition. Four of these cases were also identified as univariate outliers on either Trial 2 or 3. Due to time restrictions in the experimental sessions, six additional participants were not able to complete Trial 4. These participants were still performing Trial 4 past when the session was to have ended (i.e., past 1 hour from the start of the session), so the experimenter had to interrupt the participant, resulting in artificially low Trial 4 scores for these participants. All of these participants were in the conjunctive conditions (this is not surprising, as the instructions for the conjunctive conditions were longer than for the coactive conditions), one was in the low discrepancy condition, one was in the moderate discrepancy condition, and four were in the high discrepancy condition. These cases were also dropped from the analyses.

Dropping 13 cases from 194 left a total sample of 181 cases. The number of participants in each condition is listed in Table 4. To deal with unequal cell sizes in analyses of variance, Type IV sum of squares was used in which, "the same hypotheses are tested as in the unweighted-means approach where each cell mean is given equal weight regardless of its sample size" (Tabachnick & Fidell, 2001, p. 296).

Manipulation Checks

Discrepancy manipulation checks. Participants received feedback after Trials 1 and 2. In the questionnaire administered after the participants had completed all of the trials, participants

were asked, "Please indicate how you felt about the difference between your scores in the first individual [performance] trial and the your partner's score [the scores of person in the next room] on the first individual [performance] trial" (item numbers T10 and C10 in Appendix C). The same question was asked regarding Trial 2 (item numbers T14 and C14). Participants answered the question on the computer via moving a mouse on a scroll bar. The response scale ranged from 0 to 32767, with 0 = "scores were very similar", and 32767 = "scores were very different". The scale range was large because it was the default range for a scroll bar in the computer program. For ease of interpretation, the responses were divided by 3276.7 to create a 0 to 10 response scale. Two-way ANOVA's with performance discrepancy and work condition as the factors revealed significant main effects of performance discrepancy, for the Trial 1 discrepancy manipulation, $F(2, 175) = 182.78, p < .001$, and for the Trial 2 discrepancy manipulation, $F(2, 175) = 147.03, p < .001$. There were no main effects for work condition nor interactions between work condition and discrepancy. Bonferroni post-hoc analyses indicated significant differences in the expected direction between each pair of discrepancy levels for both the Trial 1 manipulation check and the Trial 2 manipulation check.

Participants were also asked to indicate how high or low they considered their scores and their partner's scores to be (items T11, T12, T15, T16, C11, C12, C15, and C16). Difference scores were created by subtracting the participants' ratings of their own scores from their ratings of their partners' scores. Two-way ANOVA's with performance discrepancy and work condition as the factors revealed significant main effects of performance discrepancy, for the Trial 1 difference score, $F(2, 175) = 64.00, p < .001$, and for the Trial 2 difference score, $F(2, 175) = 50.21, p < .001$. For both Trial 1 and Trial 2, Bonferroni post-hoc analyses indicated significant differences in the expected directions between the low and high discrepancy levels and between

the moderate and high discrepancy levels, but not between the low and moderate levels (though the later was in the expected direction. There were also main effects for work condition for both Trial 1, $F(1, 175) = 8.22$, $p = .005$, and for Trial 2, $F(1, 175) = 3.71$, $p = .056$ (marginally significant), such that the difference score was higher in the conjunctive conditions than in the coactive conditions. There were no interactions between work condition and performance discrepancy.

As a final discrepancy manipulation check, participants were asked in a multiple-choice question to indicate whether the difference between their scores and their partner's scores were very low, low, moderate, high, or very high (items T9, T13, C9, C13). Table 5 lists the percentage of participants that responded to each response option by condition. Participants generally responded to the question as expected such that the majority of participants in the low discrepancy condition responded that the performance difference was "low" or "very low", the majority of the participants in the moderate discrepancy condition responded that the performance difference was "moderate", and the majority of the participants in the high discrepancy condition responded that the performance difference was "high" or "very high".

Overall, the discrepancy manipulation checks suggest participants in the low, moderate, and high discrepancy conditions viewed the performance differences between them and their partner as low, moderate, and high, respectively.

Conjunctive task manipulation check. To assess whether participants in the conjunctive conditions understood how the team score was figured (i.e., that it was equal to the score of the first person to stop), participants were asked a series of multiple choice questions (items T54 through T58). In response to the question, "What happens to the team trial when one person has stopped (has three strikes)?", 89 participants (96.6%) correctly chose "It ends for both people."

One participant incorrectly responded "It continues until the other person has stopped" and two participants did not answer the question. 95.4% of the participants correctly indicated that the team score is equal to the score of the first person to stop during the team trial, three participants incorrectly answered that the team score is an average of the partner's scores, and one participant did not respond to this question. 95.4% of the participants also correctly responded that when one person stops the trial the other person can not continue, three participants incorrectly responded that when one person stops the trial the other person can continue, and one participant did not answer this question. Finally, of the 85 participants (two missing data) who answered the question, "Who was the first to stop during the first team trial?" 100% chose the response "I stopped first." Of the 86 participants (one missing data) who answered the question, "Who was the first to stop during the second team Trial?" 85 participants chose the response "I stopped first," and one participant chose the response "my partner stopped first."

Overall, the task manipulation checks suggest participants in the conjunctive conditions understood that (1) the team trial ended when the first member stopped, and (2) the team's score was equal to the score of the first member to stop.

Trial Score Corrections

The trial score for Trial 4 was the main performance measure. Trial 1 scores were retained as a baseline measure, and Trial 2 scores as a within-participant comparison to Trial 4. Trial 3 was not included in the analyses because participants had not yet experienced the conjunctive nature of the task[2].

Skewness and kurtosis values were examined for the scores from Trials 1, 2, and 4 within each of the six experimental conditions (i.e., 18 distributions were examined). All of the distributions were positively skewed. Applying logarithmic transformations (Tabachnick &

Fidell, 2001) resulted in more normal distributions for all but two of them. The transformed trial scores were thus retained for the rest of the analyses.

The individual control condition was included in the experimental design to control for fatigue effects, practice effects, and the effect of the lottery reward in the last two trials. A 2 (gender) x 2 (performance Trials 2 and 4) mixed ANOVA of the data from this condition revealed a significant main effect for trial, $F(1, 35) = 7.84$, $p = .01$, no main effect for gender, and no trial by gender interaction. Thus the observed fatigue/practice/lottery effect did not differ by gender.

To correct for effects due to fatigue, practice, and the lottery, the fourth trial score in the main (2 x 3) portion of the experimental design was multiplied by the ratio of the second to fourth trial for participants in the control condition (see Hertel et al., 2000). The mean in the control group for Trial 2 was $M = 2.1708$, and the mean in the control group for Trial 4 was $M = 2.3328$. The correction then was to multiply all other participants' fourth trial scores by $2.1708 \div 2.3328 = .9306$.

Means and standard deviations for the raw trial scores and for the transformed/corrected trial scores by condition within the main (2 x 3) portion of the experiment are listed in Tables 6 and 7. The performance hypotheses were tested using the log transformed scores for Trials 1 and 2 and the corrected log transformed scores for Trial 4.

Performance Covariates

Trial 1 was intended as a baseline measure because all participants were scored as individuals (i.e., none were scored as a team), there was no lottery involved in Trial 1, and the participants had not yet received the discrepancy manipulation. However, there was a possibility that Trial 1 performance differed by work condition, because the instructions delivered prior to

Trial 1 were different for the conjunctive and coactive conditions. Those in the conjunctive conditions participated in a small team-building activity, and team scoring was stressed during instruction (the experimenter explained the team scoring in detail and demonstrated how the team scoring worked on the computer). Further, those in the conjunctive conditions knew their performance on Trial 1 would be shared with their partner in the next room. Those in the coactive conditions also knew their performance would be shared with the other participant, but the other participant was not their teammate. In sum, those in the conjunctive conditions should have had a team mindset when entering into the first trial, whereas those in the coactive condition should not have had a team mindset. This difference in mindset may have affected the scores in Trial 1. To check for this possibility, an independent t-test between work condition was conducted. This test did not reveal a significant result ($p > .10$), so Trial 1 did not appear to differ between work conditions. Further, a two-way ANOVA with work condition and performance discrepancy as the factors did not reveal any significant effects, suggesting Trial 1 scores did not differ across experimental conditions. Trial 1 scores were significantly correlated with Trial 4 scores, $r = .42, p < .01$[3]. There was no evidence of heterogeneity of regression, suggesting the relationship between baseline performance and Trial 4 performance was the same across conditions (determined from MANOVA; Tabachnick & Fidell, 2001). There also were no significant experimenter effects. Thus, Trial 1 as a baseline measure appeared to be an appropriate covariate.

In the questionnaire administered after the completion of all four trials, participants were asked how often they play first-person shooting games such as Quake or Doom (item A53). Responses were given on a 5-point scale (1 = "never", 2 = "a few times a year", 3 = "a few times a month", 4 = "a few times a week", 5 = "every day"). Reported gaming experience was

significantly correlated with Trial 4 scores, $r = .20, p = .01$. A two-way ANOVA indicated no differences in gaming across experimental conditions. There was no evidence of heterogeneity of regression, suggesting the relationship between gaming experience and the Trial 4 performance was the same across conditions (determined from MANOVA; Tabachnick & Fidell, 2001). Gaming experience thus appeared to be an appropriate covariate.

Another possible covariate of performance on Trial 4 was the extent to which the participant valued the lottery. The three questions regarding the lottery ("How valuable is fifty dollars", "How much money is fifty dollars", "How much do you want to win fifty dollars"; items A28-A30) were all significantly correlated with one another (p's $< .01$), so they were averaged to create a "lottery score." This lottery score was not significantly correlated with Trial 4 performance. The perceived lottery value therefore was not used as a covariate of performance.

Baseline performance and gaming experience were determined to be appropriate covariates, however, they were significantly correlated with each other, $r = .27, p < .001$. This is undesirable because one covariate may make the other redundant Regression analyses determined that gaming did not account for variance in Trial 4 scores above baseline performance ($\beta = .40$, $t = 5.64$, $p < .001$ for baseline performance; $\beta = .09$, $t = 1.26$, $p = .21$ for gaming experience). When both were included in a two-way ANCOVA (with Trial 4 scores as the dependent variable and work condition and performance discrepancy as the independent variables), baseline was significantly related to Trial 4 performance, $F(1,173) = 28.22, p < .001$ and there was not a significant effect for gaming experience, $F(1, 173) = 2.17, p = .14$. Thus gaming experience appeared to be a redundant covariate when baseline performance was also included, so only baseline performance was used as a covariate of Trial 4 performance.

Hypothesis Testing

Hypothesis 1. Hypothesis 1 predicted that performance on the fourth trial would be greatest in the conjunctive condition when performance discrepancy was moderate. There were no experimenter effects for Trial 4. A two-way ANCOVA, with baseline as the covariate and work condition and performance discrepancy as the independent variables, revealed a significant main effect for performance discrepancy, $F(2, 174) = 3.84$, $p = .02$. Alpha-level adjusted for multiple comparisons is $.05 \div 3 = .017$. By this criteria, bonferroni post-hoc tests on the adjusted means revealed a marginally significant difference between the low discrepancy condition (adjusted $M = 2.05$) and the moderate discrepancy condition (adjusted $M = 2.16$), $p = .03$, and no difference between either of these and the high discrepancy condition (adjusted $M = 2.13$). Linear and quadratic contrasts were tested via a one-way ANCOVA. The linear contrast was significant, $F(1, 177) = 4.10$, $p = .04$, and the quadratic contrast was only marginally significant, $F(1, 177) = 3.37$, $p = .07$[4].

Previous research (e.g., Kerr et al., 2007) suggests motivation gain effects may be different for males and females. Thus, a three-way ANCOVA (with gender as the third factor) was conducted. The same main effect for performance discrepancy was found, $F(2, 168) = 3.60$, $p = .03$, and there were no effects of or interactions with gender.

The main performance measure is the trial score, which is equivalent to the length of the trial, because longer trials would suggest more persistence at the task. This is primarily a measure of performance quantity, how long a participant performed the task. An alternative measure is one of performance quality, how many errors the participant made while performing the task. Trials ended when participants accumulated three errors within five stimulus numbers. If a participant made an error, but then went five numbers without making another error, the

error did not count against the participant. The computer kept track of how many errors the participants made, per trial. This error count may be considered a reflection of how carefully the participant pursued the task, a measure of performance quality[5].

The error rate was calculated by dividing the number of errors by the total trial score. This measure was skewed in four of the six conditions, so a square root transformation (after adding 1 to each score) was conducted. The transformation reduced the skewness, so the transformed scores were retained for the analyses. Trial 4 error rate was significantly correlated with Trial 1 error rate (i.e., baseline error rate), $r = .18$, $p = .02$ (Trial 1 error rate was not skewed, so original values were used). Trial 4 error rate was not significantly correlated with gaming experience or perceived lottery value. A two-way ANCOVA with base error rate as the covariate and work condition and performance discrepancy as the independent variables did not reveal any significant effects. A three-way ANCOVA with gender as the third factor also did not reveal any significant effects.

Finally, a performance measure that reflected both quality of performance (error rate) and performance quantity (total score) was created by (1) creating *z*-scores for error rate and the log transformed, corrected, Trial 4 score and (2) adding those *z*-scores. This measure was slightly skewed in only two of the six conditions, so no transformations were applied. This measure was not correlated with perceived lottery value, but was significantly correlated with gaming experience, $r = .20$, $p = .01$, and with the quality/quantity measure from Trial 1 (i.e., baseline), $r = .29$, $p < .001$. The baseline measure was not skewed, so no transformations were applied. Both gaming experience and the baseline measure were significantly related to the Trial 4 quality/quantity measure in a regression analysis and in a two-way ANCOVA with discrepancy and work condition as the independent variables, so they were both retained as covariates. The

two-way ANCOVA did not reveal any significant effects of discrepancy or work condition on the Trial 4 quality/quantity measure, nor did a three-way ANCOVA with gender as the third factor.

In sum, Hypothesis 1 was partially supported such that participants in the moderate discrepancy condition had higher Trial 4 scores than those in the low discrepancy condition. However, there were no differences between the conjunctive and coactive conditions. Thus, moderate discrepancy, compared to low discrepancy, appeared to lead to higher performance in both the conjunctive and coactive conditions. Although there was not a significant performance difference between the moderate and high discrepancy conditions, the significant linear contrast suggests that higher discrepancy, compared to moderate discrepancy, may lead to higher performance. However, the mean for the high discrepancy condition was lower than the mean for the moderate discrepancy condition, and the quadratic contrast approached significance, consistent with the notion that high discrepancy may lead to *lower* performance than moderate discrepancy. Whether high discrepancy leads to higher or lower performance than moderate discrepancy, therefore, is unclear. The only conclusive result for Hypothesis 1 is that moderate discrepancy led to higher performance than did low discrepancy, across both conjunctive and coactive conditions.

Hypothesis 2. Hypothesis 2 predicted the greatest increase in performance from the second trial to the fourth trial in the conjunctive condition when performance discrepancy was moderate. This hypothesis was tested using a mixed design ANCOVA[6], with trial as the repeated measure. There was a significant main effect for trial, $F(1, 174) = 11.70, p = .001$, such that participants decreased performance from Trial 2 to Trial 4. There was also a significant interaction between trial and discrepancy, $F(2, 174) = 3.31, p = .04$. The adjusted means for

Trial 2 and Trial 4 by discrepancy condition are graphed in Figure 3. Paired t-tests were conducted separately for the low, moderate, and high discrepancy conditions. The difference between Trial 2 and Trial 4 was not significant in the moderate discrepancy condition nor in the high discrepancy condition. However, there was a significant difference between Trial 2 and Trial 4 in the low discrepancy condition.

In a three-way mixed ANCOVA with gender as the third factor, the main effect for trial remained, $F(1, 168) = 11.42, p = .001$, as did the interaction between trial and discrepancy, $F(2, 168) = 3.20, p = .04$. This analysis also revealed a marginally significant three-way interaction between trial, work condition, and gender, $F(1, 168) = 3.68, p = .06$. Explaining a three-way interaction is more straightforward using means that have not been adjusted for covariates. To determine if this was appropriate, a three-way ANOVA was run (i.e., with no covariates). The marginal three-way interaction between trial, work condition, and gender remained, $F(1, 169) = 3.71, p = .06$, so the unadjusted means were examined to explain the interaction. Two mixed design ANOVA's with trial as the repeated measure and work condition as the between factor were conducted, one using the data from the males only, and one using the data from females only. In the analysis of the males' data, there was a significant main effect for trial, $F(1, 85) = 6.23, p = .01$ such that Trial 4 scores were lower than Trial 2 scores, but no interaction between trial and work condition. For the females, there was also the significant main effect for trial, $F(1, 92) = 6.37, p = .01$, such that Trial 2 scores were lower, and there was a significant interaction between trial and work condition, $F(1, 92) = 3.94, p = .05$. Trial 2 and Trial 4 means by gender and work condition are graphed in Figure 4. It appears that, although both males and females decreased their performance from Trial 2 to Trial 4 in both work conditions, the greatest decrease was from the females in the conjunctive condition.

Error rate for Trial 2 was calculated in the same manner as the error rates for Trials 4 and 1. Error rate for Trial 2 was skewed in five of the six conditions, and a square root transformation (after adding 1 to each score) improved the skewness, so the transformed scores were retained for the analysis. A two-way mixed ANCOVA with the base error rate as the covariate, performance discrepancy and work condition as the between-subjects factors, and error rate as the repeated measure, did not reveal any significant effects. A mixed ANCOVA with gender added as the third between-subjects factor revealed a marginal three-way interaction between trial, gender, and work condition, $F(1,168) = 3.64, p = .06$. Follow-up analyses suggested the same pattern found for the total trial scores. There were no effects when just the males' data were analyzed, and there was marginal effect for the females such that they had a higher error rate on Trial 4 than on Trial 2 in the conjunctive condition, and a lower error rate on Trial 4 than on Trial 2 in the coactive condition.

Finally, a quality/quantity measure for Trial 2 was created in the same manner as that for Trials 4 and 1. This measure was skewed in only one condition, so no transformations were applied. A mixed ANCOVA with baseline quality/quantity performance and gaming experience as the covariates, discrepancy and work condition as the between-subjects factors, and the quality/quantity performance measures from Trials 2 and 4 as the repeated measure did not reveal any significant effects. A mixed ANCOVA with gender added as the third between-subjects factor also did not reveal any significant effects.

In sum, Hypothesis 2 was not supported. The hypothesis predicted an increase in performance from Trial 2 to Trial 4, but, using the scores adjusted for fatigue/practice/lottery effects, there was a decrease in performance from Trial 2 to Trial 4. Thus, a "revised" Hypothesis 2 may predict the *least* amount of decrease in the conjunctive, moderate discrepancy,

condition. Although there were no differences between work conditions, there were differences in the amount of decrease across levels of performance discrepancy. Specifically, the decrease from Trial 2 to Trial 4 was only significant in the low discrepancy condition. Further, females in the conjunctive condition exhibited the largest decrease in total score, and highest increase in error rate, suggesting these participants were the least motivated. Overall, Hypothesis 2 was not supported.

Hypothesis 3. Hypothesis 3 predicted that perceived relative ability would decrease as performance discrepancy increased. Perceived relative ability was measured by the item, "Please rate your own ability on the task compared to your partner's ability on the task [compared to the ability of the person in the next room]" (items T27 and C27). Participants answered this question on the computer via a scroll bar. Thus, perceived relative ability was measured on a continuous scale from 0 to 32767 with 0 = "my ability is much lower", and 32767 = "my ability is much higher." The scale range was large because it was the default range for a scroll bar in the computer program. For ease of interpretion, the responses were divided by 3276.7, changing the scale range to 0 (ability is much lower) to 10 (ability is much higher).

The distributions of perceived relative ability responses were examined within each condition. The distributions were normal so no transformations were necessary. There were no experimenter effects for perceived relative ability (as judged from a one-way ANOVA), and no covariates (no significant correlations with baseline performance, gaming experience, or lottery value).

A one-way ANOVA with performance discrepancy as the factor revealed a significant effect, $F(2, 178) = 7.62, p = .001$. Alpha level adjusted for multiple comparisons is $.05 \div 3 = .017$. By this criteria, bonferroni post hoc tests revealed a significant difference between low

discrepancy ($M = 4.23$, $SD = 1.86$) and high discrepancy ($M = 3.15$, $SD = 2.23$), $p = .01$, and a significant difference between moderate discrepancy ($M = 4.59$, $SD = 2.12$) and high discrepancy, $p = .001$. There was not a significant difference between low and moderate discrepancy. A two-way ANOVA with gender and performance discrepancy as the factors revealed the same main effect for discrepancy but no effect of or interaction with gender.

Hypothesis 3 was thus partially supported such that participants in the high discrepancy condition reported their ability, relative to their partner's ability, as lower than those in the low or moderate discrepancy conditions. However, inconsistent with the prediction, there were no differences between participants' ratings of their ability, relative to their partner's ability, between the low and moderate discrepancy conditions. *Hypothesis 4.* Hypothesis 4 predicted that (a) perceived indispensability would increase as performance discrepancy increased in the conjunctive conditions, and (b) perceived indispensability would not differ across discrepancy levels in the coactive conditions. In the conjunctive conditions, perceived indispensability was measured by two questions: "To what extent did the team score depend on your performance?"; and, "To what extent did the team score depend on your partner's performance?" (items T7 and T8). The response scale was the same as that for perceived relative ability (0 to 32767), so responses were divided by 3276.7, changing the range from 0 ("not at all dependent") to 10 ("very much dependent"). The second question was subtracted from the first, so a positive number indicated that the participant reported the team score was more dependent on his or her performance than on the partner's performance. The distributions of this difference score were examined within each conjunctive condition (i.e., conjunctive/low discrepancy, conjunctive/moderate discrepancy, conjunctive/high discrepancy). The distributions were normal so no transformations were necessary. There were no experimenter effects for perceived

indispensability (as judged from a one-way ANOVA), and no covariates (no significant correlations with baseline performance, gaming experience, or lottery value).

A one-way ANOVA using the data from only the conjunctive conditions did not reveal significant differences between levels of performance discrepancy. A two-way ANOVA with gender as the additional factor also did not reveal any significant effects. Thus, Hypothesis 4a was not supported.

In the coactive conditions, perceived indispensability was measured by the following two questions: "To what extent did your score depend on your performance?" and, "To what extent did the score of the person in the next room depend on your performance?" (items C7 and C8). The response scale was the same as that for indispensability in the conjunctive conditions (0 to 32767), so responses were divided by 3276.7, changing the range from 0 ("not at all") to 10 ("very much").

The intention of these questions was to parallel those used in the conjunctive conditions. A mean close to 10 was expected for the first question, and a mean close to zero was expected for the second question in the coactive conditions. The mean for the first question was $M = 3.93$, $SD = 3.40$ and the mean for the second question was $M = 7.63$, $SD = 2.46$. Although both questions are awkward, the second question is perhaps more so, and thus open to more interpretations. For example, participants may have interpreted the second question as, "To what extent was the person in the next room influenced by your score?". There was a significant correlation between the second indispensability question and perceived lottery value, $r = .26$, $p = .01$, such that participants that reported highly valuing the lottery also rated the second indispensability question higher, in the coactive conditions. The participants may have projected their desire to win the lottery onto the person in the next room, and assumed that that person too

wanted to win the lottery. Since they were told that lottery tickets were awarded based on performance, and the more lottery tickets they earned the better chance they would have of winning the lottery, perhaps those who valued the lottery were trying to get more lottery tickets than their coactor. They may have projected this onto their coactor, and assumed that they were influenced by the participant's performance because they wanted to earn more tickets than him or her. Given the awkwardness and possible interpretation problems of the second indispensability question, it was dropped from further analyses.

The distributions of responses to the first question, "To what extent did your score depend on your performance?", were examined within each coactive condition. The distributions were normal so no transformations were necessary. There were also no experimenter effects for perceived indispensability (as judged from a one-way ANOVA), and no covariates for the first question (no significant correlations with baseline performance, gaming experience, or lottery value).

A one-way ANOVA using the data from only the coactive conditions did not reveal significant differences between levels of performance discrepancy for the first indispensability question. A two-way ANOVA with gender added as a factor also did not reveal significant differences. Thus, Hypothesis 4b was supported, as it predicted that there would be no differences in indispensability across levels of performance discrepancy in the coactive conditions.

Hypothesis 5. Hypothesis five predicted that performance improvement expectancy would decrease as performance discrepancy increased. Performance improvement expectancy was measured two times - once after the first trial's discrepancy feedback, and again after the second trial's discrepancy feedback. The discrepancy manipulation entailed presenting

participants with a list of 34 scores. Their own score always ranked in the 17th place. The partner's or coactor's place in the list varied depending on the discrepancy condition (15th place in the low discrepancy condition, 11th place in the moderate discrepancy condition, and 2nd place in the high discrepancy condition). Performance improvement expectancy was measured by the question, "Assuming other participants' scores remain the same, what place do you feel you have a 100% chance of scoring in?" (items TC2 and TC5)[7]. Participants filled in a blank to indicate in which place on the list they felt they would score. This response was used as an indication of how much the participant expected his or her performance could change (i.e, increase or decrease) relative to how he or she had performed previously. Their expectancy responses were subtracted from 17, and then the difference score was divided by 17. For example, if a participant answered that he or she had a 100% chance of scoring in place 15, the expected change was an 11.76 percent increase ([(17-15)÷17] x 100). If the participant answered that he or she had a 100% chance of scoring in place 19, the expected change was a 11.76 percent decrease ([(17-19)÷17] x 100). Of main interest is the expectancy measure administered after the second trial's discrepancy feedback because (1) the participant had received the discrepancy feedback twice and (2) it is an indication of the participant's expectancy as he or she entered Trials 3 and 4, the latter providing the main performance measure of interest. Thus, the performance improvement expectancy measured after Trial 2's discrepancy feedback (item TC2) was used to test Hypothesis 5.

Ten participants answered this expectancy question with an out-of-range value. The list contained 34 places, so if the participant answered that he or she would score in a place larger than 34, this was an out-of-range response, and an indication that the participant did not understand the question. Of the ten participants who answered with an out-of-range response,

one was in the conjunctive/low discrepancy condition, three were in the conjunctive/moderate discrepancy condition, one was in the conjunctive/high discrepancy condition, two were in the coactive/moderate discrepancy condition, two were in the coactive/moderate condition, and one was in the coactive/high discrepancy condition. These cases were dropped from the following analyses.

The expectancy measure was skewed in four of six conditions, however, transforming the data did not improve the skewness, so the original values were retained for the analysis[8]. The expectancy measure was not correlated with gaming experience, or perceived lottery value, but was correlated with baseline performance, $r = -.18, p = .03$.

A two-way ANCOVA with baseline as the covariate and performance discrepancy and work condition as the independent variables did not reveal any significant differences between conditions. A three-way ANCOVA with gender as the third factor also did not reveal any significant results. Thus, Hypothesis 5 was not supported.

Hypothesis 6. Finally, Hypothesis 6 predicted that performance goals for the fourth trial would be greatest in the conjunctive moderate discrepancy condition. In the post-experimental questionnaire completed after finishing all of the performance trials, participants were asked "Did you have a performance goal for the second team trial [fourth performance trial]?" (items T40 and C40), and, "If yes, what was that goal (what score did you want)?" (items T41 and C41). Thus, not all of the participants responded to the question "If yes, what was that goal (what score did you want)?". There were 73 missing cases for this question, these were simply excluded from the analyses. Trial 4 goal was skewed within each condition, so a log transformation was applied. Further, a one-way ANOVA revealed a significant effect for experimenter, $F(3, 104) = 3.33, p = .02$. Thus, experimenter was considered a factor in the analyses. Finally, Trial 4 goal

was significantly correlated with baseline performance, $r = .43$, $p < .001$, and with gaming experience, $r = .27$, $p = .01$. However, in a two-way ANCOVA with baseline and gaming as covariates, only baseline performance was significantly related to the Trial 4 goal. Therefore, only baseline performance was retained as a covariate.

A two-way ANCOVA with baseline as the covariate did not reveal any significant effects. A three-way ANCOVA with gender also did not reveal any significant effects. When experimenter was included as a factor there was a significant main effect for work condition (in both an ANOVA without gender as a factor and in an ANOVA with gender as a factor), but this main effect was qualified by a significant interaction between work condition and experimenter. Thus, Hypothesis 6 was not supported, as no effects other than experimenter effects were found.

Model Testing

The discrepancy model was tested using multi-sample path analysis. The groups that were compared were conjunctive and coactive. The exogenous variable was the discrepancy level, and the five endogenous variables were perceived relative ability, perceived indispensability, performance improvement expectancy, performance goal, and performance. The 10 cases with out-of-range values on the performance improvement expectancy variable (see Hypothesis 5) were excluded from the path analysis.

Performance discrepancy in nature would be a continuous variable, ranging from zero (i.e., no difference in scores) to some indeterminate number greater than zero (i.e., the highest observed difference in scores). Performance discrepancy may be expressed as a percentage of the participant's score. If there was no discrepancy, then the partner's score would be 100% of the participant's score. For example, if the participant's score was 200, the partner's score would be 200 (1.00 x 200). If there was a small discrepancy such that the partner outperformed the

participant (as is the case in this study), then the partner's score would be slightly more than 100% of the participant's score. For example, is the participant's score was 200, a partner's score of 210 (1.05 x 200) would be a small discrepancy. In a correlational study, the discrepancy variable would naturally vary, with a possible range from 1.00 (partner's score equal to participant's score) to 2.00 or beyond (partner's score double, or more than double, the participant's score). Because this study was experimental, the discrepancy range was limited, from 1.05 to 1.84, and artificially restrained to be one of three values: 1.05 (low discrepancy), 1.17 (moderate discrepancy) and 1.84 (high discrepancy). Thus, there were no missing values or outliers on this variable. The variable was slightly skewed, but neither a log transformation nor a square root transformation greatly reduced the skewness, so the original values were used in the analysis. The mean discrepancy value for the conjunctive condition was $M = 1.33$ ($SD = .35$), and for the coactive condition the mean was $M = 1.36$ ($SD = .35$).

There are five endogenous variables in the discrepancy model - perceived relative ability, perceived indispensability, performance improvement expectancy, performance goal, and performance. The measure for perceived relative ability was the same as that used to test Hypothesis 3. There were no outliers and no missing cases for this variable. The perceived relative ability for the conjunctive conditions was $M = 3.50$ ($SD = 2.16$) and the mean for the coactive conditions was $M = 4.34$ ($SD = 2.07$). The measure for perceived indispensability was the first indispensability question: "To what extent did the team score depend on your performance?" for the conjunctive conditions ($M = 8.08$, $SD = 2.69$) and "To what extent did your score depend on your performance?" for the coactive conditions ($M = 7.38$, $SD = 2.49$). There were no missing values and no outliers on this variable. Although it was slightly skewed, neither a log transformation nor a square root transformation improved the skewness, so the

original values were retained for the analysis. The log transformed performance goal variable used in Hypothesis 6 was also used in the path analysis. There was one outlier, which was dropped because it was also a multivariate outlier (see below). There were 66 cases with missing values because participants were not required to report a performance goal. These cases were dealt with by instructing AMOS (statistical package used for structural equation modeling; SPSS, 2007) to estimate the means for the missing cases. The mean for the reported goal in the conjunctive condition was $M = 2.38$, $SD = .21$ and the mean for the coactive condition was $M = 2.34$, $SD = .22$. The performance measure used was the log transformed, corrected total score for Trial 4 (see Hypotheses 1 and 2). Univariate outliers and missing cases for the performance variable were dealt with before testing Hypotheses 1 and 2. The mean performance for the conjunctive condition was $M = 2.11$ ($SD = .25$) and the mean for the coactive condition was $M = 2.12$, $SD = .27$. The performance improvement expectancy measure was the same as that described under Hypothesis 5. There were four minor outliers, but they were retained for the analysis. There were 19 missing values on the expectancy variable that were dealt with by instructing AMOS to estimate the means for the missing cases. The mean expectancy (proportion increase or decrease expected) for the conjunctive condition was $M = .04$ ($SD = .28$) and the mean for the coactive condition was $M = .08$ ($SD = .24$).

There was one multivariate outlier as assessed from Mahalanobis distances (Tabachnick & Fidell, 2001), this case was also a univariate outlier on the goal variable. The case was from the coactive/low discrepancy condition and was dropped from the path analyses. There was not a collinearity problem, as assessed by tolerance and variance inflation factors. The correlations among the variables are reported for the conjunctive and coactive conditions in Tables 8 and 9.

Two models were tested. The first model tested was the theoretical model, the proposed discrepancy model illustrated in Figure 1. In the theoretical model, discrepancy was predicted to lead to perceived relative ability and performance improvement expectancy. Perceived relative ability was predicted to lead to perceived indispensability (for the conjunctive conditions only). Perceived indispensability and performance improvement expectancy were predicted to lead to performance goal, which in turn was predicted to lead to performance.

The second model tested was the experimental model, an alternative model reflecting the order of operations in the experimental procedure. This model is illustrated in Figure 5. In the experimental model, discrepancy leads to performance improvement expectancy, which then leads to performance. Performance in turn leads to perceived relative ability, perceived indispensability, and performance goal (the three variables measured after performance), and these variables co-vary.

The data reasonably fit the unconstrained theoretical model, $\chi^2(18) = 24.416$, $p = .142$, CFI = .905, RMSEA = .046. The larger the chi-square, the poorer the fit. The null hypothesis is that the data fits the model, so failure to reject the null hypothesis is support for the model. Thus, a p-value greater than .05 suggests the model is supported at the .05 level, and a p-value less that .05 suggests the model is not correct (Kline, 2005). A CFI greater than .90 suggests good fit, and an RMSEA less than or equal to .05 also suggests good fit (Kline, 2005). The analysis was also run with cross-group equality constraints imposed, $\chi^2(24) = 29.78$, $p = .192$. A chi-square difference test resulted in a non-significant difference between the constrained and unconstrained models, suggesting the factor structure is equal across groups.

The significant paths from the unconstrained theoretical model are illustrated in Figure 6. For the conjunctive group, the significant paths were from discrepancy to ability, $B = -1.453$, $\beta =$

-.232, p = .032, from ability to indispensability, B = -.276, β = -.222, p = .040, and from goal to performance, B = .922, β = .764, p <.001. For the coactive group, the significant paths were from discrepancy to ability, B = -1.972, β = -.334, p <.001 and from goal to performance, B = .658, β = .552, p <.001. It would be desirable to compute modification indices for the paths in the unconstrained model to determine whether the path coefficients are significantly different between the groups. However, these indices cannot be estimated when there is missing data. The main comparison of interest would be the path from ability to indispensability, which was significant in the conjunctive group, but not in the coactive group. The coefficients for that path in the coactive group were B = -.036, β = -.032, p = .765. The unstandardized coefficients appear to be fairly different across the two groups (B = -.276 for conjunctive, and B = -.036 for coactive).

The unconstrained experimental model was a poorer fit of the data than the unconstrained theoretical model, $\chi^2(14) = 25.165$, p = .033, CFI = .835, RMSEA = .069. For the conjunctive group, the only significant path was from performance to goal, B = .646, β = .768, p <.001. Likewise, for the coactive group, the only significant path was from performance to goal, B = .464, β = .549, p <.001. The analysis was also run with cross-group equality constraints imposed, $\chi^2(19) = 28.17$, p = .080. A chi-square difference test resulted in a non-significant difference between the constrained and unconstrained models, suggesting the factor structure underlying the experimental model is also equal across groups.

The discrepancy model was thus partially supported. Although the data fit the theoretical model reasonably well, only a few paths were significant. Performance discrepancy was negatively related to perceived relative ability as predicted. As discrepancy increased, the participant's estimate of his or her own ability, relative to his or her coactor's ability, decreased.

For the conjunctive conditions only, perceived relative ability was negatively related to perceived indispensabilty such that team members saw themselves as indispensable to the team, in a conjunctive task, when they also perceived their ability to be lower than that of the partner. Performance discrepancy did not influence performance improvement expectancy for participants in either work condition. Further, neither perceived indispensability nor performance improvement expectancy were related to performance goals in either work condition. In both work conditions, goals were significantly related to performance, but this was the case for the experimental model as well. Therefore, goals may have predicted performance, or performance may have predicted goals. Overall, the theoretical model appeared to be somewhat supported by the data.

Discussion

Study 1 examined the motivational processes of the least capable member working as a team member in a conjunctive task versus working coactively. The discrepancy model predicted that performance would be highest in the conjunctive moderate discrepancy condition. However, no performance differences by work condition were found. That is, participants in the conjunctive condition did not perform differently than those in the coactive condition. Yet, performance discrepancy did influence performance such that participants in the low discrepancy condition performed significantly worse than those in the moderate or high discrepancy condition. This effect was found in both a between-subject test (Hypothesis 1) and a mixed within/between subject test such that only participants in the low discrepancy condition had significantly lower performance from Trial 2 to Trial 4 (Hypothesis 2). Thus, moderate and high performance discrepancy between teammates or coactors appeared to be more motivating than a low performance discrepancy.

An anomalous finding was that females in the conjunctive condition appeared to be the least motivated. It may be that females in the conjunctive condition were upset that they were hurting their team's performance, and this negative affect may have reduced their performance. Another possibility is that this effect was an anomaly specific to this study, or perhaps specific to the vigilance task used in this study. Future research could explore these possibilities.

Perceived relative ability was negatively related to performance discrepancy in the model, as predicted. In the hypothesis test, it was shown that perceived relative ability was much lower when the performance discrepancy was high than when it was moderate or low. It appears that, at high discrepancy, participants attributed the difference in performance to ability rather than to effort. Relative ability was also expected to be lower when performance discrepancy was moderate, compared to when there was a low performance discrepancy. However, perceived relative ability did not significantly differ between low and moderate discrepancy conditions.

Perceived indispensability was not related to performance discrepancy in either the coactive conditions (as predicted, Hypothesis 4b), nor in the conjunctive conditions (counter to the prediction, Hypothesis 4a). However, perceived indispensability was related to perceived relative ability, for the conjunctive conditions only, as the model predicted. Thus, it appears that, indeed, participants who attributed performance differences to ability rather than effort (due to a large performance difference), also saw themselves as indispensable to the team's outcome in a conjunctive task.

Performance discrepancy did not influence performance improvement expectancy. Further, performance goal was not related to perceived indispensability nor to performance improvement expectancy. Finally, performance goal was related to performance, but this may be because the goal was measured after the performance. This suggests the first limitation of the

study - that the goal was measured after performance. This allowed the responses to the goal variable to be influenced by the performance itself, rather than the other way around. Further, had the goals been measured before performance, relationships between the goal and perceived indispensability and/or performance improvement expectancy may have been revealed. Measuring goals post-performance was thus not a good test of discrepancy model predictions regarding that variable.

Another limitation of Study 1 was the finding that performance did not differ between work conditions. Although there was a significant relationship between perceived relative ability and perceived indispensability in the conjunctive conditions, and not in the coactive conditions, this did not transfer to performance differences between work conditions.

Thus, Study 1 did not find the Kohler effect, motivation gain by the least capable member in a conjunctive task. There are a number of reasons why this study failed to find this effect. First, the Kohler effect is more pronounced under face-to-face conditions (Weber & Hertel, 2007). Participants in this study were not face-to-face, but rather were isolated from one another in separate computer rooms. Second, the Kohler effect is also more pronounced when the performance feedback is continuous, rather than given at the end of the trial (Weber & Hertel, 2007). Although participants knew whether the partner had stopped or not throughout the team trial (by the presence of the partner's ID code), a running tally of the partner's score during the team trial may have had more impact. Finally, another possibility is that the lottery incentive was a strong motivator that "trumped" any motivation that may have come from working in a team. That is, the perceived indispensability of participants in the conjunctive condition may have had no additional influence over and above their desire to win the lottery, a desire also present in the coactive condition.

The main performance difference was seen in the low discrepancy condition, where performance was lower than in either the moderate or high discrepancy conditions. Thus, a discrepancy effect did occur, just not the one that was predicted - that higher performance would occur in the moderate discrepancy condition only. Rather, higher performance was observed in *both* the moderate and high discrepancy conditions. The high discrepancy condition was perhaps not high enough. Although the manipulation checks suggested that participants viewed the discrepancy as higher in the high discrepancy condition than in the moderate discrepancy condition, the difference between the moderate and high discrepancies may not have been large enough to produce performance differences between them.

Another possibility for the discrepancy effect that was found in this study involves the expectancy variable. According to Locke and Latham (2002), "across goal levels, lower expectancies, associated with higher goal levels, are associated with higher performance" (p. 706). This would predict, then, that the least capable members should have been most motivated when discrepancy level, and its implied goal, were *high*. Consistent with this, the lowest performance occurred in low discrepancy conditions. This interpretation suggests that expectancy may have played a role in the present study. The lack of findings for the expectancy variable could be due to measurement problems. First, the set of 34 expectancy questions had a number of problems, as described. Second, the expectancy measure that was used to test the predictions asked only what place the participants felt they had a 100% chance of scoring in, not what the *highest* place was that they thought they could score in. Future research could explore the role of performance expectancy in motivation gain by employing a better measure of expectancy.

Finally, it should be noted that there were some measurement issues with perceived indispensability as well. Specifically, perceived indispensability was measured differently in the conjunctive conditions (i.e., how much the team's score depended on the participant's performance) than in the coactive conditions (i.e., how much the participant's score depended on the participant's performance). Thus, any differences in perceived indispensability between work conditions should be interpreted with some caution. As Kerr at al. (2007) noted, indispensability may be easier to manipulate than to measure, and, as they suggest, alternative designs and tests of mediation may be necessary to fully understand the role of indispensability in motivation gain.

In conclusion, performance discrepancy between members and their partners influenced the members' performance such that moderate and higher discrepancies resulted in higher performance than low discrepancy. The partial support found for the discrepancy model suggests higher discrepancies lead to higher perceived indispensability, via perceived relative ability, for those in the conjunctive conditions. However, no evidence of the Kohler effect was found, and the psychological mechanisms proposed to underlay that effect were not completely supported. Measurement and methodological problems likely contributed to the null findings - future research could test the discrepancy model with improved measurement and methodology.

Study 2: The Most Capable Member in an Additive Task

Study 2 examined the motivational processes of the most capable member as described in the discrepancy model. According to the discrepancy model, motivation should be highest for the most capable member in an additive task under conditions of moderate performance discrepancy (i.e, a social compensation discrepancy effect). Study 2 was run concurrently with Study 1, using the same general procedures as those used in Study 1.

Description of Performance Trials

Just as in Study 1, participants performed the task in four different trials, and the first trial was intended as a baseline measure of participants' ability at the task. All of the participants were scored as individuals in Trial 1, and none had yet learned of their partners' or coactor's performance. After Trial 1, participants learned that their score on Trial 1 was higher than their partner's or coactor's score on Trial 1 (i.e., the performance discrepancy manipulation). In Trial 2, the participants all continued to be scored as individuals, but under the knowledge that they had scored higher than their partner or coactor on Trial 1. After Trial 2, participants were told that their score on Trial 2 was, again, higher than their partner's or coactor's scores on Trial 2. In Trial 3, the participants in the additive conditions were scored as a team, and the team score was equal to the average of the partners' scores on that trial. The computer was programmed to lead participants to believe their partner stopped before them, and showed the team score as being the average of their scores on that trial. Importantly, that average was less than the participant's score on the trial, due to the ostensibly lower score of his or her partner. By the fourth trial, then, the participants in the additive conditions were under the impression that they had outperformed their partner on the three prior trials. Further, by the fourth trial the participants in the additive conditions had experienced the additive scoring by seeing that the team score was equal to the average of their and their partner's score, and that it was lower than their own individual score. Those in the coactive conditions were scored as individuals in Trials 3 and 4, and did not learn of their coactors' performance on those trials.

Hypotheses

Drawing from previous work (Messe et al., 2002) , Study 1 hypothesized the main performance differences to occur in Trial 4, after those in the conjunctive conditions had experienced the conjunctive nature of the task. Likewise, the main performance differences for

Study 2 were expected to occur in Trial 4, after those in the additive conditions had experienced the additive nature of the task, which is to say that the partner's lower score had brought down the team score (i.e., the team score was lower than the participant's score). The effect of work condition (additive or coactive), and the interactive effect of work condition and performance discrepancy (low, moderate, or high), on performance should thus be most evident in Trial 4. Further, these effects should also be evident in the change in performance from Trial 2 (after learning of the performance discrepancy a second time) to Trial 4 (after experiencing the additive task). The first two hypotheses reflect these predictions. Hypotheses 3 through 6 involve the motivational processes suggested by the discrepancy model as applied to the most capable member.

Hypothesis 1: Performance on the fourth trial will be greatest in the additive condition when performance discrepancy is moderate.

Hypothesis 2: The greatest increase in performance from the second trial to the fourth trial will occur in the additive condition when performance discrepancy is moderate.

Hypothesis 3: Perceived relative ability will increase as performance discrepancy increases.

Hypothesis 4a: In the additive conditions, perceived indispensability will increase as performance discrepancy increases.

Hypothesis 4b: In the coactive conditions, perceived indispensability will not vary across performance discrepancy levels.

Hypothesis 5a: In the additive conditions, performance improvement expectancy will decrease as performance discrepancy increases.

Hypothesis 5b: In the coactive conditions, performance improvement expectancy will not vary across performance discrepancy levels.

Hypothesis 6: Personal performance goals for the fourth trial will be greatest in the additive condition when performance discrepancy is moderate.

The discrepancy model will also be tested via path analysis.

Method

Participants

192 participants, 98 of whom were male, were recruited from the UIC subject pool. Of those who reported their age (n = 190), the average was 19.24 years. The study design was the same as Study 1: 2 (work condition: collective vs. coactive) x 3 (performance discrepancy: low, moderate, high) x 2 (performance trial), with the last factor repeated. With the exception of balancing gender within conditions, participants were randomly assigned to the six conditions in each study. No control condition was included in the second study because it would not have differed in any way from the control condition in Study 1. Thus, the control condition from study 1 ($N = 37$; 21 males) also served as the control condition for Study 2.

Procedure

The procedure was the same as Study 1, except (1) the team score was additive rather than conjunctive in the collective conditions and (2) the participants were always led to believe they were the most capable member rather than the least capable member.

Participants came to the lab in same sex pairs and seated at the same table, where the experimenter gave all of the verbal instructions. After completing the informed consent process, the participants in the additive condition completed the same "Getting to Know Your Partner" form as used in Study 1. The experimenter then described and demonstrated the vigilance task.

The participants then practiced the task and received information about the learning trial and the lottery. Just as in Study 1, "score" or "scores" as referred to in the procedures means the length of the trial (i.e., the total number of turns).

The participants in the additive conditions then learned how the team score would be figured. They were told that, when one partner's trial ended, the other partner would be able to continue, and the team score would be equal to the average of their scores on the team trial. The experimenter then demonstrated how the team trials work on the computer. For the coactive conditions, the experimenter demonstrated how the lottery tickets work on the computer, just as in Study 1.

After the demonstration program, participants went to their separate computer rooms and proceeded through the rest of the experiment by following the directions on the computer. As in Study 1, participants were not debriefed at the time of dismissal. Rather, they were sent a debriefing form at the end of the semester. Study 1 and Study 2 were conducted over the same semester under the same subject pool study number. Two participants were randomly drawn from the total pool of participants (i.e., all the participants from the combined studies) to receive the $50.00 reward.

The order of operations on the computer that participants encountered is the same as Study 1 (see Tables 1 and 2).

Work condition manipulation. The work condition manipulation was similar to the manipulation in the first study - it differed only in (1) the partners could continue the task when the other partner had stopped, no one was cut-off from continuing, and (2) the team score was the average of the partners' scores on the team trial, rather than the score of the first person to stop.

During the team trials, the participant's ID code (the code he or she entered into the computer before the first trial) and the partner's ID code (a fake code, every participant's "partner" had the same ID), appeared in the lower left corner. When one participant stopped the trial, his or her ID code disappeared. In the demo program shown by the experimenter, the ID codes were simply "Partner C" and "Partner D". On the first trial in the demonstration, the experimenter showed the participants what it would look like if the participant stopped before his or her partner. The experimenter performed the task for 20 numbers, then stopped responding, allowing three strikes to quickly accumulate, thereby ending the trial. The experimenter noted that her ID code (Partner C) had disappeared because she had three strikes, but that her partner's ID code (Partner D) was still visible because that person had not reached three strikes.

After the trial, the participants saw the experimenter's score (Partner C), the "partner's" score (Partner D) and the team score. Partner D's "score" was 20 more than Partner's C score. The experimenter noted that the "partner" had been able to continue even though she (the experimenter) had stopped the trial, and that the team score was equal to average of their scores on the team trial.

On the second trial in the demonstration, the experimenter showed the participants what it would look like if the partner stopped before the participant. After 15 numbers had been presented, the "partners" ID code (Partner D) disappeared. The experimenter noted that this was because the partner had three strikes, but that her ID code (Partner C) was still visible because she had not reached three strikes. The experimenter continued the task for five more numbers, then stopped responding allowing three strikes to quickly accumulate and end the trial. Again, after the trial, the participants saw the experimenter's score (Partner C), the "partner's" score (Partner D) and the team score. Partner D's "score" was 15, and Partner's C score was 23. The

experimenter noted that she had been able to continue even though the "partner" had stopped the trial, and the team score was equal to average of their scores on the team trial. During the actual trials, participants were always led to believe that their partner was the first to stop during the team trial, as the partner's ID code disappeared after a specific period of time (see discrepancy manipulation, below). The participant's score was always genuine, and the partner's ostensible score was equal to the score at which they appeared to stop the trial, to give the impression that the partner stopped first. The team score was always the average of the participant's score and the partner's ostensible score.

Just as in Study 1, for those in the coactive conditions, the participants' ID codes were not displayed during any of the trials. In the demo program, the experimenter demonstrated how the lottery tickets work, in the same way as in Study 1.

Discrepancy manipulation. The discrepancy manipulation was the same as Study 1, except participants were always led to believe that they were the most capable member rather than the least capable member.

After the first trial ("first individual trial" for those in the additive conditions, "first performance trial" for those in the coactive conditions), participants were shown a list of scores, ostensibly from past participants and from the other participants (their "partner" for those in the additive conditions, the "person in the next room" for those in the coactive conditions). The list contained 34 scores, from highest (1st place) to lowest (34th place), and the participant's score was always in the 17th place. The other scores in the list, including the partner's ostensible score, was a percentage of the participant's score (see Figure 7). In the low discrepancy conditions, the partner's score was .97 times the participant's score, and the partner's place in the list was 19 (for example, if the participant's score was 200, the partner's score was 194). In the

moderate discrepancy conditions, the partner's score was .72 times the participant's score, and the partner's place in the list was 24 (for example, if the participant's score was 200, the partner's score was 144). In the high discrepancy conditions, the partner's score was .32 times the participant's score, and the partner's place in the list was 32 (for example, if the participant's score was 200, the partner's score was 70).

After the second trial ("second individual trial" for those in the additive conditions, "second performance trial" for those in the coactive conditions), participants were shown a list of scores constructed in the same manner as that shown after the first trial. It was made clear that the first list of scores were from participants' first trial, and the second list of scores were from participants' second trial. Except for the score values in the list, which were based on the participant's second trial scores, the list was exactly the same as the list shown after the first trial (e.g., same "participants" in the same places on the list).

A pilot study (described in appendix B) confirmed that participants view the discrepancies as low, moderate, and high.

A challenge specific to Study 2 was determining at what point the participant's partner should appear to stop during the team trials. In the coactive conditions, participants were not aware of their partner's progress during the trials. In the additive conditions the partner's progress during the team trials was made available to the participants by the display of the partner's and the participant's ID codes during the trial. In the first study, participants were always led to believe that they had stopped the trial before the partner, thus the partner's ID code never disappeared. In the second study, participants needed to believe that their partner had stopped before them. The challenge, therefore, was to determine at what point during the team trial the partner's ID should disappear.

The program in the second study was such that the partner's ID disappeared when the participant reached the partner's average score from the preceding trials. For example, suppose the participant's score on the first individual trial was 200. In the moderate discrepancy condition, the partner's "score" on the first trial would have been 144 (200 times .72). Suppose the participant's score on the second individual trial was 300, the partner's "score" on the second individual trial would have been 216 (300 times .72) in the moderate discrepancy condition. On the first team trial, the partner's ID disappeared when the participant reached a score of 180. In other words, the partner's "score" on the first team trial was 180, the average of the partner's "scores" on the first two trials (144 and 216). On the second team trial, the partner's ID disappeared when the participant reached a score of 180, the average of the partner's score on the first three trials (144, 216, and 180). Thus, the partner's "score" on the second team trial was usually the same (within one point due to rounding) as on the first team trial. The intention was for the partner's ID to disappear before the participant ended the team trial, but this was not always the case, as discussed in the results section below.

Measures. The measures were exactly the same as those in Study 1 (see Appendix C for the questionnaire items).

Individual control condition. The same participants from Study 1's control condition served as controls in Study 2. These participants performed the task for four performance trials, and, like those in the coactive conditions, were scored as individuals on all four trials. However, unlike those in the coactive conditions, participants came to the lab one at a time (rather than with a coactor), and they did not get any information about other participants' performance on the task. They received the same instruction as those in the coactive conditions, except they were not given feedback about other participants' scores. Just as those in the experimental

conditions, participants in the control condition were told they would be rewarded with lottery tickets on the last two trials. Table 3 lists the order of operations on the computer for participants in the individual control condition. These participants also answered most of the questions that participants in the other conditions answered. The exceptions are noted in Appendix C.

Results

Data Cleaning for the Trial Scores

The total number of stimulus numbers presented during the trial was the trial score. This score is equivalent to the length of the trial, an indication of how long participants persisted at the task.

Trial scores were screened for univariate outliers and missing values. None of the cases were identified as univariate outliers on Trial 1, and twelve cases were identified as univariate outliers on either Trial 2, 3, or 4 (z-score > 3.29, $p < .001$, two-tailed test; Tabachnick & Fidell, 2001). None of these outliers had extreme scores on the first trial. However, one case had a z-score of 7.028 on Trial 4. This was an outlier among outliers, as most of the outliers had z-scores in a 3.50 to 5.00 range. Further, this extreme outlier was on Trial 4, the main dependent variable of interest. Therefore, it was dropped. This case was in the coactive/low discrepancy condition.

Two cases were dropped because they were missing data on one or more of the four performance trials. One case was in the additive/moderate discrepancy condition and one case was in the additive/low discrepancy condition. Three additional cases were dropped because they were prevented from finishing Trial 4 due to time restrictions. One case was in the additive/low

discrepancy condition, one was in the additive/moderate discrepancy condition, and one was in the additive/high discrepancy condition.

A challenge specific to the second study was determining at which point during the team trial the computer should indicate that the partner had stopped. The goal was to have the computer indicate the partner stopped before the participant had stopped. However, this had to vary by condition such that in the high discrepancy condition the partner stopped far before the participant was expected to stop, in the moderate discrepancy condition the partner was to stop some time before the participant stopped, and in the low discrepancy condition the partner was to stop just a bit before the participant stopped. The challenge, of course, was not knowing when the participant was going to stop. A pilot study (Appendix B) was conducted to help determine what would be the best stopping point for the partner in each discrepancy condition. However, there were still some cases for which the participant stopped before the computer indicated the partner had stopped. Specifically, 15 participants stopped before their partner "stopped" on Trial 4. Of these, 13 of them chose the response, "I stopped first" to the question, "Who was the first to stop during the second team trial?" asked in the questionnaire after the fourth trial. Two participants reported their partner stopped before them on Trial 4, but the computer records show otherwise. One of these cases (in the additive/low discrepancy condition) was also one of the cases dropped due to being prevented from finishing Trial 4. Thus, an additional 14 cases were dropped because they did not experience the additive nature of the task as intended. Not surprisingly, these cases were from the low and moderate discrepancy conditions (9 and 5, respectively).

Dropping 20 cases from 192 left a total sample of 172 cases. The sample sizes by experimental condition and gender are shown in Table 10.

Manipulation Checks

Discrepancy manipulation checks. Participants received feedback after Trial 1 and after Trial 2. In the questionnaire administered after the participants had completed all of the trials, participants were asked, "Please indicate how you felt about the difference between your scores in the first individual [performance] trial and your partner's scores [the scores of the person in the next room] on the first individual [performance] trial" (items T10 and C10 in Appendix C). The same question was asked regarding Trial 2 (items T14 and C14). Participants answered these questions on the computer via moving a mouse on a scroll bar. The response scale was a continuous scale from 0 to 32767 with 0 = "scores were very similar", and 32767 = "scores were very different". The scale range was large because it was the default range for a scroll bar in the computer program. Responses were divided by 3276.7 so the scale ranged from zero to ten.

Two-way ANOVA's revealed significant main effects of discrepancy level, for the Trial 1 discrepancy manipulation, $F(2, 166) = 134.44$, $p < .001$, and for the Trial 2 discrepancy manipulation, $F(2, 166) = 92.41$, $p < .001$. Bonferroni post-hoc analyses indicated significant differences in the expected direction between each pair of discrepancy levels. There were no main effects of work condition and no interactions between work condition and performance discrepancy.

Participants were also asked to indicate how high or low they considered their scores and their partner's scores (items T11, T12, T14, T15, C11, C12, C14, and C15). Difference scores were calculated by subtracting participants' ratings of their partners' scores from their ratings of their own scores. Two-way ANOVA's of these difference scores also revealed significant main effects for discrepancy condition: $F(2, 166) = 40.69$, $p < .001$ for the Trial 1 discrepancy manipulation, and $F(2, 166) = 23.31$, $p < .001$ for the Trial 2 discrepancy manipulation.

Bonferoni post-hoc tests showed significant differences in the expected direction between the moderate and high discrepancy conditions and between the low and high discrepancy conditions, but not between the low and moderate discrepancy conditions. There were no main effects of work condition and no interactions between work condition and performance discrepancy.

As a final check, participants were asked in a multiple-choice question to indicate whether the difference between their scores and their partner's scores were very low, low, moderate, high, or very high (items T9, T13, C9, C13). Table 11 lists the percentage of participants that responded to each response option by condition. Overall, participants appeared to view the low, moderate, and high discrepancy levels as such.

Additive task manipulation check. To assess whether participants in the additive conditions understood how the team score was figured (i.e., that it was the average of the partners' scores on that trial), participants were asked a series of multiple choice questions (items T54-T58). In response to the question, "What happens to the team trial when one person has stopped (has three strikes)?" 100% correctly responded "It continues until the other person has stopped." 100% of the participants correctly indicated that the team score is equal to the score the average of the partner's scores on that trial. 100% of the participants also correctly responded that when one person stops the trial the other person can continue. Thus, participants appeared to understand the additive nature of the team task.

Trial Scores Corrections

The trial score for Trial 4 was the main performance measure. Trial 1 scores were retained as a baseline measure, and Trial 2 scores served as a within-participant comparison to Trial 4. Trial 3 was not included in the analyses as participants had not yet experienced the

additive nature of the task[9]. There were no experimenter effects (as determined from a one-way ANOVA) for any of the trials.

Skewness and kurtosis values were examined for trial scores from Trials 1, 2, and 4 within each of the six experimental conditions (i.e., 18 distributions were examined). All of the distributions were positively skewed. Applying logarithmic transformations (Tabachnick & Fidell, 2001) resulted in more normal distributions for all but three of them. The transformed trial scores were thus retained for the rest of the analyses.

The same individual control condition used in Study 1 was used in Study 2. Therefore the same corrections were applied. To correct for effects due to fatigue, practice, and the lottery, the fourth trial scores were multiplied by the ratio of the second to fourth trial from the control condition. The mean for Trial 2 in the control condition was $M = 2.1708$, and the mean for Trial 4 in the control condition was $M = 2.3328$. The correction then was to multiply the fourth trial scores from the six experimental conditions by $2.1708 \div 2.3328 = .9306$.

Means and standard deviations for the raw trial scores and transformed/corrected trial scores by condition are listed in Tables 12 and 13. The performance hypotheses were tested using the log transformed scores for Trials 1 and 2 and the corrected log transformed scores for Trial 4.

Performance Covariates

Trial 1 was intended as a baseline measure because all participants were scored as individuals (i.e., none were scored as a team), there was no lottery involved in Trial 1, and the participants had not yet received the discrepancy manipulation. However, just as for the first study, there was a possibility in the second study that Trial 1 performance differed by work condition, because the instructions delivered prior to Trial 1 were different for the additive and

coactive conditions. To check whether a team mindset affected the scores in Trial 1, an independent samples t-test between work condition was conducted. This test did not reveal a significant result, so Trial 1 did not appear to differ between work conditions. A two-way ANOVA with work condition and performance discrepancy also did not reveal any significant differences between experimental conditions. Trial 1 was significantly correlated with Trial 4, $r = .39$, $p < .01$[10]. There was no evidence of heterogeneity of regression, suggesting the relationship between baseline performance and Trial 4 performance was the same across conditions (determined from MANOVA; Tabachnick & Fidell, 2001). Thus, Trial 1 as a baseline measure appeared to be an appropriate covariate.

Like Study 1, in the questionnaire administered after the completion of all four trials, participants were asked how often they play first-person shooting games such as Quake or Doom (item A53). They responded on a 5-point scale (1 = "never", 2 = "a few times a year", 3 = "a few times a month", 4 = "a few times a week", 5 = "every day"). Reported gaming experience was significantly correlated with Trial 4 performance, $r = .19$, $p = .01$. A two-way ANOVA indicated no differences in gaming across experimental conditions. There was no evidence of heterogeneity of regression, suggesting the relationship between gaming experience and Trial 4 performance was the same across conditions (determined from MANOVA; Tabachnick & Fidell, 2001). There were also no significant experimenter effects. Gaming experience thus appeared to be an appropriate covariate.

Another possible covariate of performance on Trial 4 was the extent to which the participant valued the lottery. Just as in Study 1, the three questions regarding the lottery ("How valuable is fifty dollars," "How much money is fifty dollars," "How much do you want to win fifty dollars;" items A28-A30) were all significantly correlated with one another (p's $< .01$), so

they were averaged to create a "lottery score". This lottery score was significantly correlated with Trial 4, $r = .19$, $p = .012$. A two-way ANOVA indicated no difference in perceived lottery value across experimental conditions. There was no evidence of heterogeneity of regression, suggesting the relationship between perceived lottery value and Trial 4 performance was the same across conditions (determined by MANOVA; Tabachnick & Fidell, 2001), and there were no experimenter effects. Perceived lottery value thus appeared to be an appropriate covariate.

Baseline performance, gaming experience, and perceived lottery value were determined to be appropriate covariates. The perceived lottery value was not correlated with baseline performance or gaming experience. However, baseline performance and gaming experience were significantly correlated with each other, $r = .17$, $p = .03$. This is undesirable because one covariate may make the other redundant Regression analyses determined that gaming accounted for some variance in Trial 4 scores above baseline performance ($\beta = .37$, $t = 5.24$, $p < .001$ for baseline performance; $\beta = .13$, $t = 1.83$, $p = .07$ for gaming experience). When both were included in a two-way ANCOVA (with Trial 4 scores as the dependent variable and work condition and performance discrepancy as the independent variables), baseline was significantly related to Trial 4, $F(1,164) = 27.43$, $p < .001$ and there was a marginally significant effect for gaming experience, $F(1, 164) = 3.21$, $p = .08$. Including both baseline performance and gaming experience as covariates reduced the error term from .059 (two-way ANOVA without covariates) to .049 (two-way ANCOVA with both covariates). When just gaming experience was used a covariate, the error term was .057, and when just baseline performance was used as a covariate, the error term was .050. Baseline performance and gaming experience appeared to be fairly redundant: So only baseline and perceived lottery value were retained as covariates.

Hypothesis Testing

Hypothesis 1. Hypothesis 1 predicted that performance on the fourth trial would be greatest in the additive condition when performance discrepancy was moderate. A two-way ANCOVA with baseline performance and lottery value as the covariates, and work condition and performance discrepancy as the independent variables, did not reveal any significant effects.

A three-way ANCOVA (with gender as the third factor) revealed only a main effect for gender, $F(1,158) = 7.31$, $p = .01$ such that males had higher scores (adjusted $M = 2.23$) than females (adjusted $M = 2.14$).

The alternative measure of performance, how many errors the participant made while performing the task, was next considered. The error rate distributions were examined within each condition. This variable was skewed in only two of six conditions, so no transformations were applied. The error rate was significantly correlated with the base error rate (number of errors in Trial 1 divided by the total score in Trial 1), $r = .42$, $p < .001$ and with gaming experience, $r = .15$, $p = .05$. Error rate was not correlated with perceived lottery value. [Note: The base error rate distributions were only slightly skewed in 3 of the conditions, so no transformations were applied]. Baseline error rate and gaming experience were not correlated with each other. Both variables accounted for variance in a regression analysis: ($\beta = .42$, $t = 6.02$, $p < .001$) for baseline error rate and ($\beta = .14$, $t = 1.95$, $p = .05$) for gaming experience. Thus, both baseline error rate and gaming experience were retained as covariates.

A two-way ANCOVA with base error rate and gaming as the covariates did not reveal any significant results for Trial 4 error rate. A three-way ANCOVA with gender as an added factor revealed a significant gender by performance discrepancy interaction, $F(1, 158) = 3.29$, $p = .04$. The adjusted means are graphed in Figure 8. Using only the data from the males, a one-way ANCOVA with performance discrepancy as the independent variable did not reveal any

significant effects. However, using only the data from females, a one-way ANCOVA did reveal a marginally significant main effect for performance discrepancy, $F(2, 84) = 2.50, p = .09$. Bonferonni post-hoc tests suggested a marginal difference between low and moderate discrepancy such that the error rate was higher in the moderate discrepancy condition, but there were no differences between moderate and high discrepancy or between low and high discrepancy.

The quality/quantity performance measure (standardized error rate plus standardized, transformed, corrected Trial 4 total score) was calculated for Trials 4 and 1 in the same manner as in Study 1. This measure was not skewed, so no transformations were applied. This measure was not correlated with perceived lottery value, but was significantly correlated with gaming experience, $r = .26, p = .001$, and with the quality/quantity measure from Trial 1 (i.e., baseline), $r = .33, p < .001$. The baseline measure was not skewed, so no transformations were applied. Both gaming experience and the baseline measure were significantly related to the Trial 4 quality/quantity measure in a regression analysis and in a two-way ANCOVA with discrepancy and work condition as the independent variables, so they were both retained as covariates. The two-way ANCOVA did not reveal any significant effects of discrepancy or work condition on the Trial 4 quality/quantity measure. A three-way ANCOVA with gender as the third factor revealed only a marginally significant main effect for gender, $F(1,158) = 5.77, p = .02$ such that males had better performance than females.

In sum, Hypothesis 1 was not supported as there were no differences across conditions in Trial 4 performance. The only significant differences were for gender such that (1) males had higher scores and better performance than females and (2) females had a lower error rate in the low discrepancy condition than in the moderate discrepancy condition.

Hypothesis 2. Hypothesis 2 predicted the greatest increase in performance from the second trial to the fourth trial in the additive condition when performance discrepancy was moderate. This hypothesis was tested using a mixed ANCOVA[11], with baseline and lottery as the covariates, performance discrepancy and work condition as between factors, and trial as the repeated measure. Only a main effect for trial was found, $F(1, 164) = 6.82, p = .01$, such that participants decreased performance from Trial 2 (adjusted $M = 2.23$) to Trial 4 (adjusted $M = 2.18$). In a mixed ANCOVA with gender as the third factor, the main effect for trial remained, $F(1, 158) = 6.23, p = .01$, but there were no other effects.

To consider change in error rate from Trial 2 to Trial 4, the error rate for Trial 2 was calculated by dividing the number of errors on Trial 2 by the total Trial 2 score. A mixed ANCOVA, with base error rate and gaming as the covariates, performance discrepancy and work condition as the between factors, and error rate from Trial 2 and Trial 4 as the repeated measure, did not reveal any significant effects. A mixed ANCOVA including gender also did not reveal any significant effects.

Finally, a quality/quantity measure for Trial 2 was created in the same manner as that for Trials 4 and 1. This measure was skewed in only one condition, so no transformations were applied. A mixed ANCOVA with baseline quality/quantity performance and gaming experience as the covariates, discrepancy and work condition as the between-subjects factors, and the quality/quantity performance measures from Trials 2 and 4 as the repeated measure did not reveal any significant effects. A mixed ANCOVA with gender added as the third between-subjects factor also did not reveal any significant effects.

In sum, Hypothesis 2 was not supported. First, the hypothesis predicted an increase in performance from Trial 2 to Trial 4, but, using the scores adjusted for practice, fatigue and

lottery effects, there was a decrease in performance from Trial 2 to Trial 4. Further, this decrease held across gender and condition. There were no conditions in which participants had *less* of a decrease, compared to decreases in other conditions.

Hypothesis 3. Hypothesis 3 predicted that perceived relative ability would increase as performance discrepancy increased. Just as in Study 1, perceived relative ability was measured with the question, "Please rate your own ability on the task compared to your partner's ability on the task [compared to the ability of the person in the next room]" (items T27 and C27) on a continuous scale from 0 to 32767 with 0 = "my ability is much lower", and 32767 = "my ability is much higher" because it was the default range for a scroll bar in the computer program. For ease of interpretation, the responses were divided by 3276.7, changing the scale range to 0 (ability is much lower) to 10 (ability is much higher).

The distributions of perceived relative ability responses were examined within each condition. The distributions were normal so no transformations were necessary. There were no experimenter effects for perceived relative ability (as judged from a one-way ANOVA). Perceived relative ability was not correlated with baseline performance, gaming experience, or perceived lottery value.

A one-way ANOVA with performance discrepancy as the factor revealed a significant effect, $F(2, 169) = 7.34, p = .001$. Alpha level adjusted for multiple comparisons is $.05 \div 3 = .017$. By this criteria, bonferroni post hoc tests revealed a marginally significant difference in the expected direction between low discrepancy ($M = 5.82, SD = 1.57$) and moderate discrepancy ($M = 6.63, SD = 1.79$), $p = .045$, and a significant difference between low discrepancy and high discrepancy ($M = 7.06, SD = 1.77$), $p = .001$. There was not a significant difference between moderate and high discrepancy. A two-way ANOVA with gender and performance discrepancy

as the factors revealed the same main effect of discrepancy, but no effects of, or interactions with, gender. Hypothesis 3 was thus partially supported such that participants in the moderate and high discrepancy conditions reported their ability, relative to their partner's ability, as higher than those in the low discrepancy conditions.

Hypothesis 4. Hypothesis 4 predicted that (a) perceived indispensability would increase as performance discrepancy increased in the additive conditions and (b) perceived indispensability would not differ across discrepancy levels in the coactive conditions. In the additive conditions, perceived indispensability was measured by two questions: "To what extent did the team score depend on your performance?"; and, "To what extent did the team score depend on your partner's performance?" (items T7 and T8). The response scale was the same as that for perceived relative ability (0 to 32767). The response scale was the same as that for perceived relative ability (0 to 32767), so responses were divided by 3276.7, changing the range from 0 ("not at all dependent") to 10 ("very much dependent"). The second question was subtracted from the first, so a positive number indicated that the participant reported the team score was more dependent on his or her performance than on the partner's performance. The distributions of this difference score were examined within each additive condition (i.e., additive/low discrepancy, additive/moderate discrepancy, additive/high discrepancy). The distributions were skewed in two of them. A square root transformation improved the skewness, so the transformed scores were used in the analysis[12]. There were no experimenter effects for perceived indispensability (as judged from a one-way ANOVA), and no covariates (no significant correlations with baseline performance, gaming experience, or lottery value).

A one-way ANOVA using the data from only the additive conditions revealed a significant effect of performance discrepancy, $F(2, 74) = 11.70$, $p < .001$. Alpha level adjusted

for multiple comparisons is .05 ÷ 3 = .017. By this criteria, Bonferroni post-hoc tests did not show a significant difference between the low discrepancy condition and the moderate discrepancy condition, $p = .176$, but did reveal significant differences in the expected directions between the low discrepancy condition and the high discrepancy condition, $p < .001$, and between the moderate and high discrepancy conditions, $p = .016$. A linear contrast was also significant, $F(1, 74) = 22.50$, $p < .001$[13]. The non-transformed means are illustrated in Figure 9. A two-way ANOVA with gender as the added factor also showed a main effect for performance discrepancy, but no gender effect or interaction with gender. Hypothesis 4a was supported such that perceived indispensability increased as performance discrepancy increased.

In the coactive conditions, perceived indispensability was measured by the following two questions: "To what extent did your score depend on your performance?"; and, "To what extent did the score of the person in the next room depend on your performance?" (items C7 and C8). The response scale was the same as that for indispensability in the additive conditions (0 to 32767), so responses were divided by 3276.7, changing the range from 0 ("not at all") to 10 ("very much").

The intention of these questions was to parallel those used in the additive conditions. A mean close to 10 was expected for the first question, and a mean close to zero was expected for the second question in the coactive conditions. The mean for the first question was $M = 7.57$, $SD = 2.44$ and the mean for the second question was $M = 2.89$, $SD = 3.35$. As noted in Study 1, both questions are awkward, but the second question is perhaps more so, and open to more interpretations. Given the awkwardness and possible interpretation problems of the second indispensability question, it was dropped from further analyses.

The distributions of responses to the first question, "To what extent did your score depend on your performance?", were examined within each coactive condition. The distributions were skewed in two conditions, but transformations (log, square rootm inverse) did not improve the skewness, so the original values were retained for the analyses. There were no experimenter effects for perceived indispensability (as judged from a one-way ANOVA), and it was not correlated with baseline performance or gaming experience, but there was a significant correlation with perceived lottery value, $r = .21$, $p = .04$.

A one-way ANCOVA with lottery as the covariate, using the data from only the coactive conditions, did not reveal significant differences between levels of performance discrepancy for the first indispensability question. A two-way ANCOVA with gender added as a factor also did not reveal significant differences. Thus, Hypothesis 4b was supported, as it predicted that there would be no differences in indispensability across levels of performance discrepancy in the coactive conditions.

Hypothesis 5. Hypothesis five predicted that performance improvement expectancy would (a) decrease as performance discrepancy increased, in the additive conditions; and (b) not be affected by performance discrepancy in the coactive conditions. Performance improvement expectancy was measured in the same way as in Study 1. The discrepancy manipulation entailed presenting participants with a list of 34 scores and their own score was always ranked in the 17th place. The partner's place in the list varied depending on the discrepancy condition (19th place in the low discrepancy condition, 24th place in the moderate discrepancy condition, and 32nd place in the high discrepancy condition). Just as in Study 1, performance improvement expectancy was measured by the question, "Assuming other participants' scores remain the same, what place do you feel you have a 100% chance of scoring in?"[14] (item TC2 and TC5).

Participants filled in a blank to indicate in which place on the list they felt they would score. This response was used as an indication of how much the participant expected his or her performance would change (i.e, increase or decrease) relative to how he or she had performed previously. Their expectancy responses were subtracted from 17, and then the difference score was divided by 17 (multiply by 100 to get a measure of percent increase or decrease the participant expected). The responses to the expectancy measure given after Trial 2's discrepancy manipulation were used to test the hypothesis because (1) participants had received the discrepancy feedback two times and (2) it should reflect their expectancy as they entered into Trials 3 and 4 (the team trials for those in the additive condition).

There were nine participants who answered this expectancy question with an out-of-range value. The list contained 34 places, so if the participant answered that he or she would score in a place larger than 34, or in place zero, this was an out-of-range response, and an indication that the participant did not understand the question. Of the nine participants who answered with an out-of-range response, two were in the additive/low discrepancy condition, one was in the additive/high discrepancy condition, three were in the coactive/moderate discrepancy condition, one was in the coactive/moderate discrepancy condition, and two were in the coactive/high discrepancy condition. These cases were dropped from the following analyses.

The expectancy measure was again skewed in four of six conditions. However, transformations (log, square root, inverse) did not improve the skewness, so the original values were used in the analysis[15]. The expectancy measure was not correlated with baseline performance, gaming experience, or perceived lottery value, and there were no experimenter effects for this variable.

A two-way ANOVA with performance discrepancy and work condition did not reveal any significant effects. A three-way ANOVA with gender as also did not reveal any significant effects. Thus, hypothesis five was not supported.

Hypothesis 6. Finally, Hypothesis 6 predicted performance goals for the fourth trial would be greatest in the additive, moderate discrepancy, condition. In the final questionnaire, after completion of all the trials, participants were asked, "Did you have a performance goal for the second team trial [fourth performance trial]?" (items T40 and C40) and "If yes, what was that goal (what score did you want)?" (items T41 and C41). Three participants reported extremely low goals, one participant reported a goal of "1", and two participants reported goals of "5". These cases were dropped from the analysis, one case was in the additive/low discrepancy condition and two cases were in the coactive/moderate condition. Because participants were not required to report a goal, there were 79 missing cases, which were not included in these analyses. This variable was skewed within three of six conditions, so a square root transformation was applied. There were no experimenter effects for Trial 4 goal. Trial 4 goal was not correlated with perceived lottery value or gaming experience, but was significantly correlated with baseline performance, $r = .41, p < .001$.

A two-way ANCOVA with baseline performance as the covariate did not reveal any significant results. A three-way ANCOVA with baseline as the covariate and gender as the third factor also did not reveal any significant effects. Thus, hypothesis six was not supported.

Model Testing

The discrepancy model was tested using a multi-sample path analysis. The groups that were compared were additive and coactive. The exogenous variable was discrepancy level. There are five endogenous variables in the discrepancy model - perceived relative ability,

perceived indispensability, performance improvement expectancy, performance goal, and performance.

As noted in Study 1, performance discrepancy in nature would be a continuous variable, and may be expressed as a percentage of the participant's score. If there was no discrepancy, then the partner's score would be 100% of the participant's score. For example, if the participant's score was 200, the partner's score would be 200 (1.00 x 200). If there was a small discrepancy such that the partner underperformed the participant, then the partner's score would be slightly less than 100% of the participant's score. For example, is the participant's score was 200, a partner's score of 194 (.97 x 200) would be a small discrepancy. In a correlational study, the discrepancy variable would naturally vary, with a possible range from 1.00 (partner's score equal to participant's score) to 0 (partner's score of zero). However, because this study was experimental, the discrepancy range was limited, from .97 to .35, and was artificially restrained to be one of three values, .97 (low discrepancy), .72 (moderate discrepancy), or .35 (high discrepancy). For the path analysis, the discrepancy amount was subtracted from 1, so that higher numbers indicated higher discrepancy. This resulted in the variable having one of three values: .03 (low discrepancy, 1 - .97), .28 (moderate discrepancy, 1 - .72) and .65 (high discrepancy, 1 - .35). There were no missing values or outliers on this variable, and the variable was not skewed. The mean discrepancy for the additive condition was $M = .37$ ($SD = .26$) and the mean for the coactive condition was $M = .33$ ($SD = .25$).

There are five endogenous variables in the discrepancy model - perceived relative ability, perceived indispensability, performance improvement expectancy, and performance. The measure for perceived relative ability was the same as that used to test Hypothesis 3. There were no outliers and no missing cases for this variable. The mean relative ability for the additive

condition was *M* = 6.67 (*SD* = 1.56) and the mean for the coactive condition was *M* = 6.41 (*SD* = 1.92). The measure for perceived indispensability was the first indispensability question: "To what extent did the team score depend on your performance?" for the additive conditions (*M* = 6.48, *SD* = 2.24) and "To what extent did your score depend on your performance?" for the coactive conditions (*M* = 7.68, *SD* = 2.34). There were no missing values and no outliers on this variable. Although it was slightly skewed, neither a log transformation nor a square root transformation improved the skewness, so the original values were retained for the analysis. The square root transformed performance goal variable used in Hypothesis 6 was also used in the path analysis. There were no outliers on this variable. There were 74 cases with missing values because participants were not required to report a performance goal. These cases were dealt with by mean estimation in the AMOS program (SPSS, 2007). The mean for the reported goal in the additive condition was *M* = 15.25 (*SD* = 4.61) and the mean in the coactive condition was *M* = 14.88 (*SD* = 5.41). The performance measure used was the log transformed, corrected total score for Trial 4 (see Hypotheses 1 and 2). Univariate outliers and missing cases for the performance variable were dealt with before testing Hypotheses 1 and 2. The mean performance for the additive condition was *M* = 2.18 (*SD* = .22) and the mean for the coactive condition was *M* = 2.19 (*SD* = .26). The performance improvement expectancy measure was the same as that described under Hypothesis 5. There were no outliers on this variable. There were 25 missing values, which dealt with my mean estimation in AMOS (SPSS, 2007). The mean expectancy (proportion increase or decrease) for the additive condition was *M* = .08 (*SD* = .26) and the mean for the coactive condition was *M* = .03 (*SD* = .26).

There were no multivariate outliers as assessed by Mahalanobis distances. There was not a collinearity problem, as assessed by tolerance and variance inflation factors. The correlations among the variables by work condition are listed in Tables 14 and 15.

Just as for Study 1, two models were tested. The first model tested was the theoretical model, the proposed discrepancy model. The second model tested was the experimental model, an alternative model reflecting the order of operations in the experimental procedure. These are the same models tested in Study 1 (i.e., same variables and paths).

The data only somewhat fit the unconstrained theoretical model, $\chi^2(18) = 27.921$, p = .063, CFI = .793, RMSEA = .059. The larger the chi-square, the poorer the fit. The null hypothesis is that the data fits the model, so failure to reject the null hypothesis is support for the model. Thus, a p-value greater than .05 suggests the model is supported at the .05 level, and a p-value less that .05 suggests the model is not correct (Kline, 2005). A CFI greater than .90 suggests good fit, and an RMSEA less than or equal to .05 suggests good fit Kline, 2005). The analysis was also run with cross-group equality constraints imposed, $\chi^2(24) = 41.52$, $p = .015$. A chi-square difference test resulted in a significant difference between the constrained and unconstrained models at a .05 level, suggesting the factor structure is not equal across groups.

The significant paths from the unconstrained theoretical model are illustrated in Figure 10. For the additive groups, the significant paths were from discrepancy to ability, $B = 1.624$, $\beta = .26$, $p = .019$, from ability to indispensability, $B = .746$, $\beta = .519$, $p < .001$, and a marginally significant path from goal to performance, $B = .013$, $\beta = .264$, $p = .073$. For the coactive group, the significant paths were from discrepancy to ability, $B = 2.071$, $\beta = .269$, $p = .009$, from ability to indispensability, $B = .252$, $\beta = .207$, $p = .047$, from indispensability to goal, $B = -1.162$, $\beta =$ -

.511, $p < .001$, and a marginally significant path from goal to performance, $B = .010$, $\beta = .211$, $p = 097$.

The data did not fit the unconstrained experimental model, $\chi^2(14) = 31.63$, $p = .005$, CFI = .633, RMSEA = .088. For the additive groups, the path from performance to goal was marginally significant, $B = 5.338$, $\beta = .257$, $p = .089$. For the coactive group, the significant paths were from performance to ability, $B = 1.323$, $\beta = .178$, $p = .091$ (marginally significant) and from performance to goal, $B = 6.305$, $\beta = .307$, $p = .017$. The analysis was also run with cross-group equality constraints imposed, $\chi^2(19) = 33.12$, $p = .023$. A chi-square difference test resulted in a non-significant difference between the constrained and unconstrained models, suggesting the factor structure is equal across groups.

The discrepancy model appears to be only somewhat supported in the second study, as the data moderately fit the theoretical model. Performance discrepancy was positively related to perceived relative ability as predicted. As discrepancy increased, the participant's estimate of his or her own ability, relative to his or her coactor's or partner's ability, also increased. In both the additive and coactive conditions, perceived relative ability was positively related to perceived indispensabilty such that team members saw themselves as indispensable when they also perceived their ability to be higher than that of the partner. The model predicted this would occur under additive conditions, as the stronger member should recognize that the team's performance relies heavily on his or her performance, and that dependency increases as the discrepancy between the stronger member and the weaker member widens. The finding that perceived relative ability influenced perceived indispensability in the coactive conditions was inconsistent with the model's predictions. Performance discrepancy did not influence performance improvement expectancy for participants in either work condition. Further, performance

improvement expectancy was not related to performance goals in either work condition. In the coactive condition only, perceived indispensability was negatively related to goal, such that goals were lower when perceived indispensability (how much the participant's score depended on his or her performance) was higher. This finding is inconsistent with the model, as the model predicted a positive relationship between indispensability and personal goal. In both work conditions, goals were marginally significantly related to performance, but this was the case for the experimental model as well. Therefore, goals may have predicted performance, or performance may have predicted goals. Overall, the theoretical model appeared to be somewhat supported by the data, although some unexpected relationships were revealed in the coactive condition.

Besides differences in the path coefficients between the groups, there was a difference in the overall factor structure between the additive and coactive conditions, for the theoretical model. This suggests that the model does not apply across groups, that the groups are not "equal". In other words, the relationships proposed in the model may apply to one group, but not to the other.

Comparison of Conjunctive and Additive Conditions

The theoretical and experimental models were also tested in a multi-sample analysis comparing the conjunctive group from Study 1 and the additive group from Study 2. The variables used in Study 1 were identical to those used in Study 2, except for the transformed goal variable. In study one, this variable was log transformed, and in study two this variable was square-root transformed. The square root transformation improved the skewness in Study 1, whereas the log transformation increased the skewness in Study 2. Thus, the square root

transformed goal variable was used in these analyses. The mean for the square-root transformed goal for the conjunctive condition was $M = 15.87$ ($SD = 3.91$).

This data somewhat fit the unconstrained theoretical model, $\chi^2(18) = 27.47$, $p = .071$, CFI $= .872$, RMSEA $= .058$. The analysis was also run with cross-group equality constraints imposed, $\chi^2(24) = 80.21$, $p < .001$. A chi-square difference test resulted in a significant difference between the constrained and unconstrained models at a .001 level, suggesting the factor structure is not equal across groups. The significant paths from the unconstrained model are the same as reported above, except for the path between goal and performance in the conjunctive condition. The estimate for that path remained significant at a $p = .01$ level, but changed from $\beta = .768$ to $\beta = .755$ ($B = .050$).

The data did not fit the unconstrained experimental model, , $\chi^2(14) = 26.308$, p = .024, CFI = .833, RMSEA = .076. The analysis was also run with cross-group equality constraints imposed, $\chi^2(19) = 31.69$, $p = .034$. A chi-square difference test resulted in a non-significant difference between the constrained and unconstrained models, suggesting the factor structure was equal across groups for the experimental model.

Overall, the discrepancy model appears to offer some explanation for the motivation of both the least capable member in a conjunctive task *and* the most capable member in an additive task. In both cases, performance discrepancy influenced perceived relative ability, such that both members appeared to attribute performance differences to differences in ability, rather than effort, as the discrepancy widened. Also, both the least capable member and the most capable member appeared to consider their relative abilities when judging how indispensable they were to their team's outcome. For the least capable member in a conjunctive task, this meant that he or she saw that the less able he or she was compared to the partner, the more his or her team's

performance depended on his or her individual performance. Likewise for the most capable member in an additive task, the more able he or she was in comparison to the partner, the more his or her team's performance was dependent on his or her individual performance. Unfortunately, the members' perceptions of their indispensability to the team did not influence their performance goals or their performance. Further, performance improvement expectancy was not related to performance discrepancy, goals, or performance, for either member. Finally, and importantly, there was a difference in the overall factor structure between the additive and conjunctive conditions, for the theoretical model. This suggests that the model does not apply to both groups, that the proposed processes do not explain the motivation of both the least capable member in a conjunctive task and the most capable member in an additive task.

Discussion

Study 2 examined the motivational processes of the most capable member working as a team member in an additive task versus working coactively. The discrepancy model predicted that performance would be highest in the additive moderate discrepancy condition. However, no performance differences by either work condition or discrepancy level were found. An anomalous finding was that females in the low discrepancy condition had a lower error rate, suggesting they were more careful in their performance. Perhaps these participants wanted to keep their status as the higher performer, and thus paid more attention to the task. However, it is unclear why this would occur only for the females. Future research could explore whether this effect was driven by status concerns.

Perceived relative ability was positively related to performance discrepancy in the model, as predicted. In the hypothesis test, it was shown that perceived relative ability was lower when the performance discrepancy was low compared to when the performance discrepancy was

moderate or high. It appears that, as discrepancy increased, participants attributed the difference in performance to ability.

Perceived indispensability was related to performance discrepancy in the additive conditions (as predicted, Hypothesis 4a), but not in the coactive conditions, (as predicted, Hypothesis 4b). However, perceived indispensability was related to perceived relative ability in both the additive and coactive conditions. Members who attributed performance differences to ability rather than effort (due to a large performance difference), also saw themselves as indispensable to the team's outcome in an additive task. There was also a positive relationship between ability and indispensability in the coactive conditions. In these conditions, the indispensability question was, "To what extent did your score depend on your performance?". Perhaps, when the participant felt that he or she was able at the task, he or she was more likely to attribute good outcomes (i.e., scores) to his or her performance than to some other, unknown source (e.g., computer error). That is, perhaps these participants were more likely to make internal attributions than external attributions.

Tolli and Schmidt (2008) examined the effect of attributions and feedback on self-efficacy and performance goals. They found that participants who made internal attributions and were given positive performance feedback reported higher self-efficacy and higher performance goals than those who made external attributions. Indeed, for those in the coactive conditions in this study, there was a significant positive correlation between ability and performance improvement expectancy. These same participants also may have made an internal attribution, as suggested by the positive relationship between perceived relative ability and their response to the question, "To what extent did your score depend on your performance?". Thus, the performance discrepancy feedback indicating they performed better on the task than their partner

(i.e., positive feedback), may have led them to believe that they were good at the task (ability), were likely to be able to improve their performance (expectancy), and their scores were a function of their performance and not some outside influence (indispensability or internal attribution). It is possible that this occurred only in the coactive conditions because, under additive conditions, the performance discrepancy suggests a need for compensating for the partner. This concern for the most capable member in an additive task may have negated the positive effects of positive feedback that was seen in the coactive condition.

Although positive feedback may have led to internal attributions and higher performance expectancy, consistent with Tolli and Schmidt (2008), positive feedback in this study did not lead to higher performance goals. In fact, for the coactive participants, higher levels of perceived indispensability (or internal attributions) were associated with lower performance goals, a finding that is inconsistent with Tolli and Schmidt (2008). However, in their study, Tolli and Schmidt (2008) informed the participants that they had performed either above or below some non-social, or absolute, benchmark. In the current study, participants were told they had scored above their partner, allowing for a downward social comparison that was not present in the Tolli and Schmidt study. Perhaps this downward comparison, in absence of any concern for compensating for a lower performing partner, was ultimately de-motivating, and although it led to more internal attributions (i.e., higher indispensability), it also led to lower performance goals.

Finally, performance improvement expectancy was not related to any other variables for those in the additive conditions, and performance goals were moderately related to performance in both the additive and coactive conditions. As noted in Study 1, though, the goal was measured after performance. This allowed the responses to the goal variable to be influenced by the

performance itself, rather than the other way around. Therefore, measuring goals post-performance was not a good test of the discrepancy model predictions regarding that variable.

The results of Study 2 showed that performance did not differ between work conditions or levels of performance discrepancy. Thus, Study 2 did not find the social compensation effect, motivation gain by the most capable member in an additive task. As noted in Study 1, participants were isolated from one another in separate computer rooms. Perhaps the social compensation effect, like the Kohler effects, is more pronounced in face-to-face conditions. Similarly, performance feedback was given at the end of the trial rather than continuously during the team trial. Again, perhaps the social compensation effect, like the Kohler effect, is more likely when the performance feedback is continuous. Another possibility that was also mentioned in Study 1 is that the lottery incentive may have been a strong motivator that "trumped" any motivation that may have come from working in a team. That is, the perceived indispensability of participants in the additive condition may have had no additional influence over and above their desire to win the lottery, a desire also present in the coactive condition.

Study 2 had an additional consideration that was not present in Study 1, namely the most capable member's concern that he or she was being taken advantage of by the least capable member. This is akin to "the sucker effect" (Kerr, 1983), that occurs when team members exert less effort because they believe their co-members are also exerting less effort. Participants in the present study appeared to attribute larger performance differences to relative abilities, but lower discrepancies may have suggested differences due to effort rather than ability. If so, participants in the lower discrepancy conditions would be less likely to exert more effort in order to avoid being the "sucker". However, performance did not differ across levels of performance discrepancy, suggesting those in the moderate and high discrepancy conditions may have also

been concerned about being the "sucker". Although the results of the present study suggested participants attributed their partner's lower performance (particularly at higher levels of discrepancy) to their partner's lower ability, relative to their own, the extent to which they attributed the partner's lower performance to effort was not measured. Including a measure of perceived relative effort would have allowed a clearer understanding of the participant's perceptions of their partner's lower performance, and, perhaps, whether the participant was concerned with being the sucker.

The finding that the factor structure underlying the theoretical model was different across the additive and coactive groups suggests that there are different motivational processes for these groups. The performance discrepancy between members may have a lot more meaning to the most capable member in an additive task than to the most capable member in a coactive task. In an additive task, working with a lower performer may elicit concerns about compensating for the lower performer and/or the concern of being the sucker. In a coactive task the performance discrepancy simply supplies the most capable member with a downward social comparison - it should not elicit the same concerns that performance discrepancy in an additive task does. Further, the underlying factor structure may have been different because, perhaps, the indispensability measure in the additive condition was measuring a different construct than that in the coactive condition. As noted, the latter may have been tapping attributions. Better measurement could help clarify some differences between the additive and coactive groups. More likely, though, the most capable member under these different conditions simply experiences different motivational processes.

Finally, the same measurement problems that were found in Study 1 were also present in Study 2. The set of 34 expectancy questions had a number of problems, and the expectancy

measure that was used to test the predictions asked only what place the participants felt they had a 100% chance of scoring in, not what the *highest* place was that they thought they could score in. Also, perceived indispensability was measured differently in the additive conditions (i.e., how much the team's score depended on the participant's performance) than in the coactive conditions (i.e., how much the participant's score depended on the participant's performance).

In conclusion, performance discrepancy between the participant and the partner did not influence the participant's performance. Partial support for the discrepancy model suggests higher discrepancies lead to higher perceived indispensability, via perceived relative ability, for those in the additive conditions. For those in the coactive conditions, higher relative ability may influence both their performance attributions (perceived indispensability) and their performance improvement expectancy. However, these effects were not a part of the proposed discrepancy model. No evidence of the social compensation effect was found, and the psychological mechanisms proposed to underlie that effect were not completely supported. Measurement and methodological problems likely contributed to the null and inconsistent findings. Future research could test the discrepancy model with improved measurement and methodology, and may consider revising the model to include attributions and a relationship between relative ability and performance improvement expectancy. Serious consideration should also be given to the notion that the processes undergone by the most capable member in an additive task simply are different from the processes undergone by the most capable member in a coactive task.

General Discussion

In an effort to explain why some members work hard in their teams while others exert little effort, this paper proposed and tested a model that suggested the amount of performance discrepancy between members influences their motivation to work harder in their teams than

they do alone, a motivation gain. The two studies investigated the psychological processes behind motivation gain by (a) the least capable member in a conjunctive task (the Kohler effect) and (b) the most capable member in an additive task (the social compensation effect). Although these motivation gains appeared on the surface to parallel one another, the extent of their similarity had not yet been empirically evaluated. To the author's knowledge, this project was the first to examine both types of motivation gain within the same paradigm. So, did these studies suggest that the Kohler effect and the social compensation effect are "two sides of the same motivational coin" (Williams, et al., 2003, p. 340)? The answer is, probably not. They probably are two different coins altogether.

The underlying factor structure was shown to be different across the conjunctive and additive groups. That is, the processes for the least capable member in a conjunctive task are likely different from the processes for the most capable member in an additive task. There are several reasons why these processes may differ. First, the nature of the performance discrepancy feedback is different. For the least capable member, the feedback provides an upward social comparison, which may be motivating under both conjunctive and coactive conditions. For the most capable member, the performance discrepancy feedback provides a downward social comparison, which may be de-motivating in both the additive and coactive conditions. Second, the least capable member in a conjunctive task may want to improve performance to *avoid hurting* the team (e.g., an avoidance motivation), whereas the most capable member in an additive task may want to compensate to *help* the team (e.g., an approach motivation). Further, the most capable member has the additional concern of being the sucker, a concern in direct contradiction to a concern for helping the team. These contradicting concerns apply only to the most capable member in the additive task, the least capable member in a conjunctive task should

not experience this contradiction. Third, the nature of the conjunctive and additive tasks are fundamentally different. In the conjunctive task, the team's performance is completely influenced by the performance of the least capable member, whereas in the additive task the team's performance is only partially dependent on the most capable member (because the least capable member also contributes to the team outcome). Thus, there are several reasons why the motivational processes of the least capable member in a conjunctive task may be fundamentally different from the motivational processes of the most capable member in an additive task.

Although the processes of the least capable member in a conjunctive task and the most capable member in an additive task may be two different coins, one finding in the present studies suggested that, at least, they are both coins. The amount of performance discrepancy between members influenced their perceived relative ability, which in turn influenced their perceived indispensability to their teams. The least capable member in a conjunctive task saw himself or herself as the less able member when performance discrepancy was moderate to high, and in turn saw himself or herself as indispensable to the team. Similarly, the most capable member in an additive task saw himself or herself as the more able member when performance discrepancy was moderate to high, and in turn saw himself or herself as indispensable to the team. This suggests that indispensability may be an underlying motivator in *both* the Kohler effect and the social compensation effect. However, in these studies no relationship was found between perceived indispensability and performance, so the actual role played by indispensability is unclear at best.

This paper began with quotes regarding teamwork. Now another colloquialism is offered: "Don't throw the baby out with the bathwater." Although not supported in the present studies, the discrepancy model does suggest a number of insights and future avenues for

understanding motivation gain. First, there were a number of methodological and measurement concerns that could be addressed in future work and thus provide a better test of the model. Second, a discrepancy effect for the least capable member was found such that a low discrepancy led to lower performance than a moderate or high discrepancy. Third, although it did not transfer to performance differences, a type of discrepancy effect was also found for the most capable member in an additive task such that a high performance discrepancy led to higher perceived indispensability. This suggests that a social compensation discrepancy effect may exist, which has not previously been proposed. Finally, examining the model as applied to the least capable and most capable members under coactive conditions suggests a number of ways in which the model may be improved.

Some of the methodological and measurement constraints have already been mentioned - participants were not face-to-face, they did not receive continuous feedback during the team trials, the lottery incentive may have "trumped" any possible additional motivation from perceived indispensability of the teammates over the motivation of the coactors, goals should have been measured before performance, and there were problems with the measures for performance improvement expectancy and perceived indispensability. The last issue, regarding perceived indispensability, is perhaps the least easily corrected. Perceived indispensability may be easily measured in team conditions (e.g., "To what extent did your team score depend on your performance? On your partner's performance?"), but measuring a similar construct in coactive conditions proved challenging. The present studies aimed to compare perceived indispensability of team members with perceived indispensability of coactors, but the questions in the coactive conditions seemed not to have been interpreted as intended ("To what extent did your score depend on your performance?", "To what extent did the score of the person in the next room

depend on your performance?"). The second question appeared to be interpreted as "To what extent was the person in the next room influenced by your performance?", and, as discussed in Study 2, the second question may have been tapping attributions rather than indispensability. That is, it may have been tapping the extent to which participants attributed their scores to internal sources (i.e., their own performance) than to external sources (e.g., computer errors). This dilemma of measuring indispensability may not be easily resolved. As Kerr et al. (2007) suggest, indispensability may be more easily manipulated than measured.

Another methodological concern not previously discussed is the way in which performance feedback (i.e. the discrepancy manipulation) was administered. Participants were given a list of 34 scores, in which they were told they scored in the 17th place. Thus, all participants were told they were "average" at the task, but that their partners were above or below average. The intention was to hold normative feedback constant while varying the feedback relative to the partner's. However, the participant may not have been as concerned with the partner's performance as he or she would have been had that been the *only* information he or she received. For example, the participants may have been more concerned with being "average" than with their standing relative to their partner. Further, in setting a performance goal, the participants may have chosen a place higher on the list as a goal, ignoring completely the information about their partner or coactor. Future research could easily eliminate this problem by providing only feedback regarding the partner's performance and the participant's performance.

A final methodological concern involves the vigilance task that was used in these studies. Increased motivation and effort translate to increased performance on this task, as evidenced not only in the first pilot study (Appendix A) but also in the control condition in which participants increased their performance on the last two trials, the trials in which lottery tickets were awarded

(admittedly, this may have been a practice effect, but it also may have been due to being motivated to win the lottery). Thus, although the vigilance task appears to be an appropriate task for testing motivation gain, perhaps a more conservative approach would be to test of the discrepancy model using a task in which motivation gain has been previously demonstrated (e.g., a physical persistence task).

The Kohler discrepancy effect found in previous studies (Messe et al., 2002), that motivation gain was higher under moderate discrepancy than under low or high discrepancy, was not completely replicated in Study 1. Although Study 1 found lower performance under low discrepancy, there were no significant differences between moderate and high discrepancies. One explanation for this may be methodological - perhaps the high discrepancy was not high enough. Had the discrepancy in the high discrepancy condition been larger, perhaps the Kohler discrepancy effect may have been demonstrated. The means trended in the expected direction (an inverted-U shape between discrepancy and performance), and a quadratic contrast approached significance. The way in which discrepancy was communicated, in the list of 34 scores with the participant's score in the middle, may not have illustrated enough "distance" in the list between the participant and the partner. The other explanation for not demonstrating the Kohler discrepancy effect is theoretical. Perhaps, as Locke and Latham's (2002) work suggests, higher goals, implied by higher discrepancies, lead to higher performance. This suggests that the discrepancy model may be revised with a closer look at goal theory.

The discrepancy model suggests that performance discrepancy implies a level of performance improvement and thus influences performance improvement expectancy of reaching a higher level of performance. This expectancy, in turn, was predicted to lead to performance via the adoption of personal performance goals. The body of literature examining

goals likely has implications for the discrepancy model. Some of this literature also examines the role of self-efficacy or performance expectancy in goal setting and goal acceptance, as well as the role of positive and negative feedback (e.g., Bandura & Cervone, 1983; Tolli & Schmidt, 2008). Thus, future work could explore different ways of measuring different types of goals (e.g., implied goals and accepted goals) as well as different ways of measuring performance improvement expectancy. This work may also consider performance discrepancy as positive or negative feedback, and the impact of different kinds of feedback on expectancy and goal setting (e.g., Bandura & Cervone, 1983; Tolli & Schmidt, 2008).

The studies presented here assumed motivation would be manifested in performance and in stated goals. Other manifestations of motivation could be examined, such as self-report of effort, desire to perform well, or perceived need to perform well. However, these are still just assumed manifestations of motivation. Motivation, like many other psychological concepts, is a latent variable that can only be inferred from other variables. How accurately any one variable captures motivation is debatable, and future research would do well to include several dependent variables that may reflect motivation, rather than assuming motivation will be reflected in performance or goals alone.

Examining the processes in the discrepancy model under coactive conditions can also provide important insight. For example, the finding in Study 2 that perceived relative ability was related to perceived indispensability in the coactive conditions prompted more careful thought about the meaning of the indispensability variable in those conditions. As discussed, the question intended to measure indispensability in the coactive conditions may have actually been measuring attributions. This suggested another way, then, in which the discrepancy model may be improved - attributions may be an important moderator or mediator of performance

discrepancy and expectancy. Future work could explore how attributions may influence motivation in both a coactive and collective setting. Examining the coactive conditions also prompted more careful thought about the different processes the least capable member and the most capable member may undergo. The underlying factor structure was equal across the coactive and conjunctive conditions for the least capable member perhaps because this member experiences similar processes (e.g., upward social comparison) under both conditions, the only difference being an additional "boost" of indispensability in the conjunctive condition. On the other hand, the underlying factor structure was not equal across the coactive and additive conditions for the most capable member, suggesting this member experiences different processes under the different conditions. Although in both conditions the most capable member makes a downward comparison, that comparison may be more meaningful when the comparison other is a teammate than when he or she is a coactor. When the partner is a teammate, the most capable member is faced with either helping the team by compensating or, in contrast, avoiding being the sucker.

Regardless of the challenges involved in doing so, examining the underlying processes of motivation gain should involve a comparison with coactive conditions. Without such a comparison, the "value added" of being in a team setting is unknown. Performance discrepancy between coactors influences their performance (e.g., Seta, 1982), and it is only by comparing those processes to that which takes place under collective conditions that researchers can begin to understand when and why some members work harder in their teams relative to working alone.

As suggested by the finding that performance discrepancy influenced perceived indispensability of the most capable member in an additive task, a social compensation

discrepancy effect may exist. However, in these studies indispensability did not lead to performance, so a social compensation effect was not demonstrated. This may be for a number of methodological reasons already mentioned. Further, the participants may have been concerned with being the sucker, if they felt their partner was not exerting effort on the task. Finally, both a social compensation discrepancy effect and the Kohler discrepancy effect may simply be difficult to demonstrate in a laboratory setting.

With a few exceptions (e.g., Liden et al., 2004) the general Kohler effect and the social compensation effect have been demonstrated in laboratory studies. In these studies, researchers were able to manipulate the feedback to the participants that led them to believe they were either the least capable member or the most capable member. The researchers were able to determine the necessary level of ability discrepancy between members needed to produce gains in motivation by the members. Demonstrating a discrepancy effect, in which multiple levels of discrepancy are communicated, likely requires further fine-tuning of the manipulated discrepancy levels. Such fine-tuned levels of ability discrepancy could be specific to features of the particular study, and thus replication and extension of discrepancy effects may prove difficult to demonstrate across laboratory studies. This does not negate the existence of the effects, however. There are a number of features of real-world teams that are difficult to imitate in the laboratory, but that may be important for eliciting a motivation gain by team members.

First, members of actual teams may be concerned about their reputation. Members may work harder in their team because they want to maintain a good performance record, both within their teams and across teams. That is, a kind of impression-management tactic may be to work harder in the team than when working alone. This could occur particularly when others outside the team evaluate a team member by the team outcome. For example, the outcome of a research

team can have a large impact on any one member's chances of securing future research funding or positions.

A second feature of real teams is the members' knowledge of their fellow members' relative abilities. In the laboratory, this knowledge is communicated by the experimenter or by a confederate, and the participant learns of the discrepancy only shortly before entering the team task. In the real-world, team members have existing, and perhaps well-established, knowledge of team members' abilities. The knowledge of ability discrepancy in actual teams is thus likely richer and more extensive than in the ad-hoc teams in the laboratory. Indeed, even before Kohler's original studies began the ability discrepancies between his participants, members of a rowing club, were likely well-established and well-known within the club.

Third, most teams in actual settings have a past and a future. Past interactions likely contribute to the knowledge of fellow members' capabilities, as well as their personalities, work ethics, and communication styles. Future interactions may influence members' motivation in their teams due to the desire to be accepted as a valuable member and the need to work together in a congenial manner. Fourth, ability differences between members vary naturally in actual teams, perhaps providing a larger range of ability discrepancy. As noted, one difficulty in demonstrating influences of varying levels of ability discrepancy in the laboratory is establishing appropriate differences between low and moderate discrepancy and between moderate and high discrepancy. Further, ability differences that vary naturally form a continuum whereas ability differences manipulated in the laboratory are restricted to specific levels, compromising the continuous nature of the variable. Another feature of real teams that may influence members' motivation is that team performance is usually multi-faceted, the team outcome does not rely on

one specific ability across members. Thus, team members weak in one area may help the team by increasing performance in another area.

Finally, actual team members likely place more importance on the task, and on the team itself, than laboratory participants. In Kohler's original studies, the participants were members of a rowing club. Their team (the rowing club) may have been important to them, and they may have been motivated to perform well in front of, or for, their fellow teammates. Further, the weight-lifting task they performed may have been important to them. Being athletes, physical ability may have held some personal importance. This personal importance of the task and the team may be less likely in ad-hoc teams in the laboratory. Thus, some of the restrictions inherent in controlled laboratory studies may make discrepancy effects difficult to demonstrate and/or replicate in those settings. Studies of actual teams may have more success in demonstrating an influence of varying levels of ability discrepancies between members and their motivation to perform in the team.

This suggests that future directions for this line of work may consider several features of actual teams that are typically not present in the laboratory. The tasks and teams should be of high importance to the member, because the member's reputation is at stake and/or because the team or the task is personally important to the member. Task importance may be heightened in the laboratory by "matching" the participants to the task. For example, participants who define themselves as "athlete" may place more importance on a physical task than participants who do not consider "athlete" important to their self-concepts. Another way in which the importance of the task may be heightened is by linking the team outcome to the member in a more personal way. For example, a quasi-experimental design may compare two models of college student teams - one in which the team outcome is included in a job resume, and one in which the team

outcome is not included in a resume. Assessing teams in which the ability discrepancy between members varies naturally, and assessing enough teams to obtain a large range of discrepancies, are also important for future directions. Further, the possibility of helping the team with increasing performance in alternative areas could be examined if the team outcome was multi-faceted, requiring a diverse array of member abilities.

Team members can, and do, work harder in their teams than working alone. This gain in performance has been demonstrated primarily in laboratory studies. This literature has shown that both the least capable member and the most capable member can be motivated to increase performance in their teams relative to working alone. However, demonstrating and replicating the impact of various levels of ability discrepancy may be more difficult in the laboratory. Examination of teams and tasks in which the team members have a high investment may be more likely to reveal the impact of the ability discrepancies between members and their subsequent motivation to perform well in their teams.

References

Bandura, A., & Cervone, D. (1983). Self-evaluative and self-efficacy mechanisms governing the motivational effects of goal systems. *Journal of Personality and Social Psychology, 45*, 1017-1028.

Baranski, J. V., Thompson, M. M., Lichacz, F. M., McCann, C., Gil, V., Pasto, L., & Pigeau, R. A. (2007). Effects of sleep loss on team decision making: Motivational loss or motivational gain? *Human Factors, 49,* 646-660.

Bevan, W., & Turner, E. D. (1965). Vigilance performance with a qualitative shift in reinforcers. *Journal of Experimental Psychology, 70*, 83-86.

Bruning, J. L., & Kintz, B. L. (1977). *Computational handbook of statistics.* Glenview, IL: Scott, Foresman, and Company.

Delaney, H. D., & Maxwell, S. E. (1981). On using analysis of covariance in repeated measures designs. *Multivariate Behavioral Research, 16,* 105-123.

Earley, P. C., & Lituchy, T. R. (1991). Delineating goal and efficacy effects: A test of 3 models. *Journal of Applied Psychology, 76*, 81-98.

Geister, S., Konradt, U., & Hertel, G. (2006). Effects of process feedback on motivation, satisfaction, and performance in virtual teams. *Small Group Research, 37*, 459-489.

Hardt, J. W., Bridgett, D. J., & Karau, S. J. (2001). Coworker ability and effort as determinants of individual effort on a collective task. *Group Dynamics: Theory, Research, and Practice, 3*, 181-190.

Helton, W. S., Dember, W. N., Warm, J. S., Matthews, G. (1999). Optimism, pessimism, and false failure feedback: Effects on vigilance performance. *Current Psychology, 18.*

Hertel, G., Deter, C., & Konradt, U. (2003). Motivation gains in computer-supported groups. *Journal of Applied Social Psychology, 33,* 2080-2105.

Hertel, G., Kerr, N. L., & Messe, L. A. (2000). Motivation gains in performance groups: Paradigmatic and theoretical developments on the Kohler effect. *Journal of Personality and Social Psychology, 79*, 580-601.

Hertel, G., Niemeyer, G., & Clauss, A. (2008). Social indispensability or social comparison: The why and when of motivation gains of inferior group members. *Journal of Applied Social Psychology, 38,* 1329-1363.

Huey, B. M., & Wickens, C. D. (1993). *Workload transition: Implications for individual and team performance.* Washington, DC: National Academy Press.

Karau, S. J., & Williams, K. D. (1993). Social loafing: A meta-analytic review and theoretical integration. *Journal of Personality and Social Psychology, 65,* 681-706.

Karau, S. J., & Williams, K. D. (1997). The effects of group cohesiveness on social loafing and social compensation. *Group Dynamics: Theory, Research, and Practice, 1*, 156-168.

Kerr, N. L. (1983). Motivation losses in small groups: A social dilemma analysis. *Journal of Personality and Social Psychology, 45,* 819-828.

Kerr, N. L., Messe, L., Park, E. S., & Sambolec, E. J. (2005). Identifiability, performance feedback, and the Kohler effect. *Group Processes and Intergroup Relations, 8*, 375-390.

Kerr, N. L., Messe, L. A., Seok, D., Sambolec, E., Lount, R., & Park, E. S. (2006, May). Task demands and the Kohler effect. In N. L. Kerr & G. Hertel (Chairs), *Motivation gains of inferior group members*. Symposium conducted at the meeting of the Midwestern Psychological Association, Chicago, IL.

Kerr, N. L., Messe, L. A., Seok, D., Sambolec, E., Lount, R., & Park, E. S. (2007). Psychological mechanisms underlying the Kohler motivation gain. *Personality and Social Psychology Bulletin, 33,* 828-841.

Kline, R. B. (2005). *Principles and practice of structural equation modeling.* New York: The Guilford Press.

Liden, R. C., Wayne, S. J., Jaworski, R. A., Bennett, N. (2004). Social loafing: A field investigation. *Journal of Management, 30,* 285-304.

Locke, E. A. & Latham, G. P. (2002). Building a practical useful theory of goal setting and task motivation. *American Psychologist, 57*, 705-717.

Lount, R. B., & Phillips, K. W. (2007). Working harder with the out-group: The impact of social category diversity on motivation gains. *Organizational Behavior and Human Decision Processes, 103,* 214-224.

Manly, T., Robertson, I. H., Galloway, M., & Hawkins, K. (1999). The absent mind: Further investigation of sustained attention to response. *Neuropsychologia, 37*, 661-670.

Messe, L. A., Hertel, G., Kerr, N. L., Lount, R. B., & Park, E. S. (2002). Knowledge of partner's ability as a moderator of group motivation gains: An exploration of the Kohler discrepancy effect. *Journal of Personality and Social Psychology, 82*, 935-946.

Robertson, I. H., Manly, T., Andrade, J., Baddeley, B. T., & Yiend, J.(1997). 'Oops!': Performance correlates of everyday attentional failures in traumatic brain injured and normal subjects. *Neuropsychologia, 35*, 747-758.

SPSS. (2007). AMOS 16.0. *Statistical Package for the Social Sciences.* Chicago, IL.

Sambolec, E. J., Kerr, N. L., & Messe, L. A. (2007). The role of competitiveness at social tasks: Can indirect cues enhance performance? *Journal of Applied Sport Psychology, 19,* 160-172.

Seta, J. J. (1982). The impact of comparison processes on coactor's task performance. *Journal of Personality and Social Psychology, 42,* 281-291.

Seta, J. J., Seta, C. E., & Donaldson, S. (1991). The impact of comparison processes on coactors' frustration and willingness to expend effort. *Personality and Social Psychology Bulletin, 17,* 560-568.

Shepperd, J. A. (1993). Productivity loss in performance groups: A motivation analysis. *Psychological Bulletin*, 113, 67-81.

Shepperd, J. A. (1995). Remedying motivation and productivity loss in collective settings. *Current Directions in Psychological Science, 4,* 131-134.

Shepperd, J. A., & Taylor, K. M. (1999). Social loafing and expectancy-value theory. *Personality and Social Psychology Bulletin, 25*, 1147-1158.

Smallwood, J., Davies, J. B., Heim, D., Finnigan, F., Sudberry, M., O'Connor, R., & Obonsawin, M. (2004). Subjective experience and the attentional lapse: Task engagement and disengagement during sustained attention. *Consciousness and Cognition, 13,* 657-690.

Smith, R. L., Lucaccina, L. F., & Epstein, M. H. (1967). Effects of m1tary rewards and punishments on vigilance performance. *Journal of Applied Psychology, 51*, 411-416.

Steiner, I. D. (1972). *Group process and productivity.* New York: Academic Press.

Stroebe, W., Diehl, M., & Abakoumin, G. (1996). Social compensation and the Kohler effect: Toward a theoretical explanation of motivation gains in group productivity. In E. H.

Witte & J.H. Davis (Eds.), *Understanding Group Behavior* (pp. 37-65). Mahwah, NJ: Erlbaum.

Tabachnick, B. G. & Fidell, L. S. (2001). *Using multivariate statistics.* Boston: Allyn and Bacon.

Tolli, A. P. & Schmidt, A. M. (2008). The role of feedback, causal attribution, and self-efficacy in goal revision. *Journal of Applied Psychology, 93,* 692-701.

Vroom, V. H. (1964). *Work and motivation.* New York: Wiley.

Vrugt, A., & Koenis, S. (2002). Perceived self-efficacy, personal goals, social comparison, and scientific productivity. *Applied Psychology: An International Review, 51,* 593-607.

Vrugt, A., Oort, F. J., & Zeeburg, C. (2002). Goal orientations, perceived self-efficacy and study results amongst beginners and advanced students. *British Journal of Educational Psychology, 72,* 385-397.

Warm, J. S., Kanfer, F. H., Kuwada, S., & Clark, J. L. (1972). Motivation in vigilance: Effects of self-evaluation and experimenter-controlled feedback. *Journal of Experimental Psychology, 92*, 123-127.

Weber, B., & Hertel, G. (2007). Motivation gains of inferior group members: A meta-analytical review. *Journal of Personality and Social Psychology, 93,* 973-993.

Williams, K. D., Harkins, S. G., & Karau, S. J. (2003). Social performance. In M. A. Hogg & J. Cooper (Eds.), *The Sage Handbook of Social Psychology* (pp. 327-346). Thousand Oaks, CA: Sage Publications.

Williams, K. D., & Karau, S. J. (1991). Social loafing and social compensation: The effects of expectations of co-worker performance. *Journal of Personality and Social Psychology, 61*, 570-581.

Witte, E. H. (1989). Kohler rediscovered: The anti-Ringlemann effect. *European Journal of Social Psychology, 19*, 147-154.

www.caseystengel.com. July 30, 2008.

www.psyc.bbk.ac.uk/research/DNL/stats/Repeated_Measures_ANCOVA. June 5, 2008.

www.vincelombardi.com. July 30, 2008.

Endnotes

1. Although participants were told the lottery tickets would be awarded based on their performance, in actuality two participants were randomly chosen to receive the reward.

2. Although not included in the main analyses, a number of exploratory analyses were conducted for Trial 3 scores. The score distributions across the six conditions were found to be non-normal, and a log transformation was applied to Trial 3 scores. Further, a practice/fatigue/lottery correction was also applied to Trial 3 in the same manner as applied to Trial 4: the third trial scores in the main 2x3 portion of the experiment were multiplied by the ratio of the second to third trial for those in the control condition (see Hertel et al., 2000). The mean for Trial 2 was $M = 2.1708$, and the mean for Trial 3 was $M = 2.2687$. The correction then was to multiply the third trial score by $2.1708 \div 2.2687 = .9568$. Trial 3 scores were correlated with baseline performance and gaming experience, but gaming experience appeared to be a redundant covariate. A three-way ANCOVA with baseline as the covariate, and experimenter, work condition, and performance discrepancy as the independent variables did not find any significant effects. This may be attributed to the fact that participants had not yet experienced the conjunctive nature of the task. That is, Trial 3 was the first trial in which they were scored as a team, and the impact of team scoring may not be evident until Trial 4. This interpretation is consistent with the Messe et al. (2002, Study 2) study in which a significant discrepancy effect was found only in Trial 4 of four, not in Trial 3 of four. Thus, Trial 4 scores were used as the main dependent variable.

3. The correlation between the log transformed Trial 1 and Trial 4 scores ($r = .42$, $p < .01$) was different than the correlation between the raw Trial 1 and Trial 4 score ($r = .50$, $p < .01$). The correlation between the log transformed Trial 1 score and the raw Trial 4 score was $r = .46$, $p <$

.01, and the correlation between the raw Trial 1 score and the log transformed Trial 4 score was $r = .43, p < .01$. Trial 1 raw scores were also submitted to an independent t-test between work condition, and did not differ by condition ($t < 1.00$). The transformed Trial 1 scores were used in the analyses of covariance.

4. Contrasts assume equal intervals between the levels of the variable. The difference between the discrepancy conditions in response to the manipulation check question, "Please indicate how you felt about the difference between your score in the first performance trial and the score of your partner/the person in the next room on the first performance trial" were examined. The response scale was a continuous scale ranging from 0 to 32767 with 0 = "scores were very similar", and 32767 = "scores were very different". Two correlations between this question and the condition numbers were calculated. In the first correlation, the condition number was either 1 (low discrepancy condition) or 2 (moderate discrepancy). The correlation between condition number and reported discrepancy was $r = .59, p < .001$ such that reported discrepancy increased as condition number (i.e., manipulated performance discrepancy) increased (from low to moderate discrepancy). In the second correlation, the condition number was either 2 (moderate discrepancy condition) or 3 (high discrepancy). The correlation between condition number and reported discrepancy was $r = .69, p < .001$ such that reported discrepancy increased as condition number (i.e., manipulated performance discrepancy) increased (from moderate to high discrepancy). A test for the significance of the difference between correlations was conducted following the procedures outlined in Bruning and Kintz (1977). It was concluded that the two correlations were not significantly different from one another, suggesting that the increase in perceived discrepancy from low to moderate discrepancy was not different from the increase in

perceived discrepancy from moderate to high discrepancy. Thus, the assumption of equal intervals in linear and quadratic contrasts appeared to be met.

5. Hit rate (number of hits divided by total trial score) is another measure of performance quality. However, the hit rate is simply one minus the error rate, and so the two are perfectly negatively correlated.

6. The following information was retrieved on June 5, 2008 from www.psyc.bbk.ac.uk/research/DNL/stats/Repeated_Measures_ANCOVA. According to this source, the main effect of a repeated measure may disappear when adding a covariate because the covariance of the covariate with the difference score is used to adjust the sum of squares. This can be corrected by mean centering the covariate (Delaney & Maxwell, 1981, cited as source for mean centering correction). To center the baseline performance covariate, each individual's baseline score was subtracted from the overall mean baseline score ($M = 2.1784$). Likewise, to center the base error rate, the error rates were subtracted from the overall error rate mean ($M = .058$), and to center the gaming variable, each gaming score was subtracted from the mean gaming score ($M = 2.10$).

7. The expectancy questionnaire included three questions: (1) "You have achieved a score in the 17th place. Assuming other participants' scores remain the same, what is the chance that you will score in that place again? Please indicate your answers as a percentage, ranging from 0% chance to 100% chance. ___%" (items TC1 and TC 4 in Appendix C); (2) "Assuming other participants' scores remain the same, what place do you feel you have a 100% chance of scoring in? ____ place" (items TC2 and TC5); and (3) "Considering your answer to question 2, please indicate the chance you have of scoring in each of the following places (again, assuming other participants' scores remain the same). ____ % for 1-34 places" (items TC3 and TC6). The last question was

actually a set of 34 questions. The responses to the last set of 34 questions did not correspond to the answers to the first two questions. For example, if a participant had responded to the first question that he or she had a 100% chance of scoring in place 17 again, the response to the 17th question in the set of 34 questions should have been 100%. However, only 42% of the participants responded in that manner. That is, 58% of the participants responded differently to the first question than to question 17 of the set of 34. It was expected that most participants would respond to these two questions with exactly the same response, so that the correlation between the two would be very high. However, the correlation was $r = .582, p < .001$, indicating rather low reliability between these two items. The participants' responses to the second question also did not correspond to their responses to the set of 34 questions. For example, if a participant had responded to the second question that he or she had a 100% chance of scoring in place 15, the response to the 15th question in the set of 34 questions should have been 100%. However, only 25% of the participants responded in that manner. That is, 75% of the participants responded differently to the second question than to the corresponding question in the of the set of 34. Only 13% of participants responded to the first, second, and third questions as expected. Finally, an examination of the responses to the set of 34 questions suggested other problems. For example, several participants responded by putting the same number in every place, (e.g., responding that he or she had a 10% chance of scoring in place 1, 2, 3, and so forth to place 34), or did not complete the entire measure, leaving numerous missing values (e.g., only responding to the questions for first and last place). Given these numerous problems, the responses to the set of 34 questions were not used in the analyses.

8. Because the transformation were applied to proportion scores, the scores were converted to all positive numbers greater than one by adding two to each score. Log, square root, and inverse transformations did not improve the skewness.

9. Although not included in the main analyses, a number of exploratory analyses were conducted for Trial 3 scores. Just as for Study 1, the score distributions across the six conditions were found to be non-normal, and a log transformation was applied to Trial 3 scores. Further, like study 1, a practice/fatigue/lottery correction was also applied to Trial 3 in the same manner as applied to Trial 4: the third trial scores in the main portion of the experiment were multiplied by the ratio of the second to third trial from those in the control condition. The mean for Trial 2 was $M = 2.1708$, and the mean for Trial 3 was $M = 2.2687$. The correction then was to multiply the third Trial score by ($2.1708 \div 2.2687 = .9568$). There were no experimenter effects for Trial 3. Trial 3 scores were correlated with baseline performance. A two-way ANCOVA with baseline as the covariate, and work condition and performance discrepancy as the independent variables did not find any significant effects. This may be attributed to the fact that participants had not yet experienced the additive nature of the task. That is, Trial 3 was the first trial in which they were scored as a team, and the impact of team scoring may not be evident until Trial 4. This interpretation is consistent with the Messe et al. (2002, Study 2) study in which a significant Kohler discrepancy effect was found only in Trial 4 of four, not in Trial 3 of four. Thus, Trial 4 scores were used as the main dependent variable.

10. The correlation between the log transformed Trial 1 and Trial 4 scores ($r = .39, p < .01$) was different than the correlation between the raw Trial 1 and Trial 4 score ($r = .33, p < .01$). The correlation between the log transformed Trial 1 score and the raw Trial 4 score was $r = .34, p < .01$, and the correlation between the raw Trial 1 score and the log transformed Trial 4 score was r

= .36, $p < .01$. Trial 1 raw scores were also submitted to an independent samples t-test between work condition, and did not differ by condition. The transformed Trial 1 scores were used in the analyses of covariance.

11. Just as in Study 1, the covariates used in the mixed-ANCOVA's were mean centered. To center the baseline performance covariate, each individual's baseline score was subtracted from the overall mean baseline score ($M = 2.1844$). To center the lottery covariate, each individual's lottery score was subtracted from the overall mean lottery score, $M = 20256.13$. For the analyses using error rate, to center the gaming covariate, each individual's gaming score was subtracted from the overall mean gaming score, $M = 2.13$, and to center the base error rate, each individual's base error rate lottery score was subtracted from the overall mean base error rate, $M = .0591$.

12. Before applying the square root transformation, 11 was added to the indispensability scores to make them all greater than one.

13. Contrasts assume equal intervals between the levels of the variable. The differences between the discrepancy conditions in response to the indispensability measure were examined. Two correlations between this question and the condition numbers were calculated. In the first correlation, the condition number was either 1 (low discrepancy condition) or 2 (moderate discrepancy). The correlation between condition number and indispensability was $r = .35$, $p = .02$ such that indispensability increased as condition number (i.e., manipulated performance discrepancy) increased (from low to moderate discrepancy). In the second correlation, the condition number was either 2 (moderate discrepancy condition) or 3 (high discrepancy). The correlation between condition number and indispensability was $r = .33$, $p = .01$ such that indispensability increased as condition number (i.e., manipulated performance discrepancy) increased (from moderate to high discrepancy). A test for the significance of the difference

between correlations was conducted following the procedures outlined in Bruning and Kintz (1977). It was concluded that the two correlations were not significantly different from one another, suggesting that the increase in perceived indispensability from low to moderate discrepancy was not different from the increase in perceived indispensability from moderate to high discrepancy. Thus, the assumption of equal intervals in linear and quadratic contrasts appeared to be met.

14. The same general problems found with the expectancy measure in Study 1 were also found in Study 2.

15. Because transformations were applied to the proportional score, the scores were converted to all positive numbers greater than one by adding two to each score (so a score of 2.00 meant no change, above 2.00 meant expected increase, and below 2.00 meant expected decrease).

Appendix A

Pilot Study 1: Appropriateness of a Visual Vigilance Task for Studying Motivation Gain

This report describes a visual vigilance task developed for the purpose of studying motivation gain. Vigilance tasks involve a participant recognizing and responding to target stimuli amongst non-target stimuli. Early research suggested that performance on vigilance tasks can be influenced by motivation (Bevan & Turner, 1965; Smith, Lucaccina, & Epstein, 1967; Warm, Kanfer, Kuwada, & Clark, 1972), therefore such a task may be appropriate for studying motivation gain. Any task used to study performance effects of motivation should be one in which there is a monotonic relationship between effort and performance; increased effort should be reflected in increased performance (e.g., Hertel, et al., 2003). This pilot study was conducted to assure this was the case for the vigilance task.

Key Features of the Task

Motivation gain is often studied using a physical task, for example, holding a weight above a trip rod for as long as possible (Hertel, et al., 2000; Messe, et al., 2002; Stroebe, et al., 1996). In this type of task, the motivated participant holds the weight longer than the unmotivated participant. That is, if motivated, the participant exerts more effort and persists at the task, even after becoming fatigued. In developing a vigilance task to study motivation gain, the goal was to emulate the features of the physical task used in past research, substituting cognitive fatigue for physical fatigue. Specifically, the vigilance task required the participant to sustain attention, concentrate, as long as possible. The motivated participant should persist and sustain attention longer than the unmotivated participant, and this should result in performance differences. Two key features were incorporated into the development of the new vigilance task to assure the task required participants' concentration.

The first key feature is the "response" of a non-response such that participants were to press the space bar when the target and stimuli did not match, and withhold any response when the target and stimuli did match. The second key feature is the participants must keep in memory the given target to which they compare the presented number. The given target is a target number *previously*, not currently, displayed on the computer screen. That is, participants have to compare the stimuli to a target number that was previously displayed. Details of the task are reported under Study 1.

The only difference between the task used in this pilot study and the one used in the main studies is that, in this pilot study, when participants viewed their trial scores at the end of the trial, they saw the total number of turns (presentations), the number of hits (correct responses) and the number of misses and false alarms (incorrect responses), as well as a review of what constitutes a hit, a miss, and a false alarm (these definitions are also given as part of initial instruction). In the main studies, the misses and false alarms were not distinguished, rather they were simply referred to as "errors".

Pilot Study Overview

All of the participants in the pilot study performed: (1) a short practice trial; (2) a learning trial; (3) a first performance trial; and (4) a second performance trial. All participants received "low effort" instruction on first performance trial. Half of the participants also received "low effort" instruction on the second performance trial. The other half received "high effort" instruction on the second performance trial. It was expected that participants in the high effort condition would perform significantly better on the second performance Trial than those in the low effort condition.

Method

Participants. 24 students from the UIC Psychology Subject Pool participated. 15 were male and the average age was 18.67 years.

Procedure. Participants were brought to the lab in pairs, only because it was more effective to go over the task instruction with two people at a time. After the task instruction, participants were moved to separate rooms and proceeded independent of one another. Before moving the participants to their separate computer rooms, the experimenter described the task and the scoring. The experimenter then had the participants watch her perform the practice trial. Ten numbers were presented in the practice trial, regardless of the number of errors (i.e., it was possible to have ten errors). Participants then moved to their separate computer stations and performed the practice trials by themselves. From this point on the experimenter gave instruction to the participants individually.

After the practice trial, the experimenter introduced the learning trial. She explained that the trial would be the same as the practice trial, only (1) the targets would be different, (2) the trial would continue until 3 strikes accumulated, (3) the numbers would be presented increasingly faster, (4) a maximum of 100 numbers would be presented during the practice trial (note, in the main studies the learning trial maximum was 50 numbers) and (5) a "Help" key was available on the first trial.

For the learning trial only ("Trial 1" to the participants), the participants were able to recall the previous target by pressing the "Enter" key. When this key was pressed, the previous target would briefly appear. It was explained that this help was being provided on the first trial only, because they were still learning the task. The participants were told they could use that help key as often as they wanted, but to try not to use it too much, because it would not be available on later trials. Finally, the participants were instructed to view their scores when the

trial ended, take a little break, and to get the experimenter (waiting in an adjacent room) when they were ready to proceed to Trial 2.

Before the first performance trial ("Trial 2" to the participants), the experimenter reminded participants that the "Help" key would no longer be available. They were then instructed that, for the second trial, they should "Stop the trial as soon as you BEGIN to START feeling tired". To stop the trial, they were to simply stop responding, which allows 3 strikes to accumulate very quickly and end the trial. Again, participants were told that when the trial ended, they should look at their scores for the Trial, take a break, and get the experimenter when ready to proceed to Trial 3.

Before the second performance trial ("Trial 3" to participants), the experimenter informed the participants that this would be the last trial, and instructed them to write their scores for the trial on a small slip of paper (given to them) at the end of the trial. The purpose of this was simply because Trial 3 scores would be filled in on a later questionnaire. Then, half the participants (the low effort condition) were given the same instruction as the previous trial (i.e., to stop when tired). The other half of the participants (the high effort condition) were told "This time, go as long as you can. Even when you start feeling tired, proceed for as long as possible". They were told the reason for this was because the task may be similar to tasks companies use to hire employees, as it may be a good measure of how long one can concentrate even after becoming fatigued. They were asked to imagine they themselves were applying for a job with great benefits and hours, and to get the job they had to do their ABSOLUTE best on the task.

Once participants completed the last trial, they were given a questionnaire. The questionnaire included self-reports of effort for Trials 1, 2, and 3, ratings of enjoyment of the task, the difficulty of the task, and a space for comments. Gender and age were also asked.

[Note: Participants also answered the question, "What would scores would another student have to get on Trial 3 for you to believe that student's ability was different from your own?". These results are not reported here, but were used to plan Pilot Study 2.] Effort, enjoyment, and difficulty were rated on a 0 to 9 scale with 0 = very little and 9 = very much. After the questionnaire the participants were debriefed, thanked, and dismissed.

Results

The main dependent variable was the total number of stimuli that were presented in each performance trial. A mixed factorial ANOVA, with performance trial total as the within-subject factor and condition as the between-subject factor, yielded a significant trial by condition interaction. Participants in the high effort condition increased their total score from performance Trial 1 ($M = 153.00$, $SD = 70.40$) to performance Trial 2 ($M = 248.67$, $SD = 115.98$), whereas participants in the low effort condition did not increase their total score from performance Trial 1 ($M = 165.75$, $SD = 81.98$) to performance Trial 2 ($M = 165.50$, $SD = 103.39$), $F(1,22) = 9.212$, $p = .006$. A between-subjects ANOVA with the total from performance Trial 2 as the dependent variable yielded consistent results. There was a significant main effect for condition such that participants in the high effort condition had a significantly higher total than participants in the low effort condition, $F(1,22) = 3.438$, $p = .077$. Finally, the amount of change in total numbers presented from performance Trial 1 to 2 was calculated by dividing Trial 1 scores by Trial 2 scores, such that a number greater than one indicated an increase in performance, and a number less than one indicated a decrease in performance. A between-subjects ANOVA found a main effect of condition such that participants in the high effort condition increased their performance from Trial 1 to Trial 2 ($M = 1.80$, $SD = .83$) and participants in the low effort condition decreased their performance, ($M = .98$, $SD = .31$), $F(1,22) = 10.400$, $p = .004$. These three

analyses were also run for total number of hits per performance trial (total number of correct responses) as the dependent variable and the findings paralleled those of the analyses using total number of presentations as the dependent variable. No gender effects were found for the total score or for the number of hits.

Participants' self-reports of effort were consistent with their performance scores. A mixed factorial ANOVA, with reports of effort for performance Trials 1 and 2 as the within-subject factor and condition as the between-subject factor, yielded a significant trial by condition interaction. Participants in the high effort condition reported increasing their effort from performance Trial 1 ($M = 6.50$, $SD = 2.714$) to performance Trial 2 ($M = 8.42$, $SD = .996$), whereas participants in the low effort condition reported little increase in effort from performance Trial 1 ($M = 7.33$, $SD = 1.557$) to performance Trial 2 ($M = 7.75$, $SD = 1.960$), $F(1,22) = 3.816$, $p = .064$.

Finally, there were no differences in gender or condition on ratings of task difficulty or enjoyment. Participants generally found the task to be challenging, but not impossible. The mean difficulty rating was 4.79 on a 0-9 scale, where 0 = not at all difficult and 9 = very difficult ($SD = 2.48$). Participants reported the task as enjoyable as well. The mean enjoyment rating was 6.54 on a 0-9 scale where 0 = not at all enjoyable and 9 = very enjoyable ($SD = 2.06$).

Discussion

The purpose of this pilot study was to determine the appropriateness of a new visual vigilance task for studying motivation gain. The new task incorporated features intended to require the participant to sustain attention. The pilot test was conducted to assure that increased effort on the new task would be reflected in increased performance. Indeed, participants who were told to increase effort also had increased performance, compared to participants who were

not told to increase effort. The newly developed visual vigilance task thus appeared to be appropriate for studying motivation effects on performance.

Appendix B

Pilot Study 2: Manipulation Checks for Performance Discrepancy and Team Task

This report describes a pilot study conducted to assure participants viewed the manipulated discrepancy levels as intended (i.e., as low, moderate, or high), and to make sure they understood the nature of the team task (i.e., conjunctive or additive).

Pilot Study Overview

All of the participants in the pilot study performed the visual vigilance task described in Study 1. They performed six trials: (1) a short practice trial; (2) a learning trial; (3) a trial in which they were scored as individuals, not as a team; (4) a second individual trial; (5) a trial in which they were scored as a team; and (6) a second team trial. The procedures were identical to those described in Studies 1 and 2, except for the placement of the manipulation check questions. In the main studies, manipulation checks occurred after the completion of the last trial. In this pilot study, manipulation checks for discrepancy levels occurred immediately after receiving the performance feedback/list of scores, which occurred after each individual trial. The task manipulation for the task occurred before and after each team trial. Participants in this pilot study did not complete an expectancy measure.

Method

Participants. 102 students from the UIC Psychology Subject Pool participated. 45 were male and the average age was 18.82 years.

Procedure. The procedures were identical to those described in the main studies for participants in the team conditions.

Measures. The measures were similar to those described in the main studies for participants in the team conditions (Appendix C). The only exception was the expectancy

measure was not included in the pilot study, and manipulation checks occurred between the trials rather than after the last trial.

Results

Discrepancy Check for Conjunctive Conditions. 54 students were in the conjunctive conditions: 16 in the low discrepancy condition, 18 in the moderate discrepancy condition, and 20 in the high discrepancy condition.

As the first manipulation check, participants were instructed as follows: "Using the scroll bar below, please indicate how you felt about the difference between your scores in the first individual trial and your partner's scores on the first [second] individual trial. Using your mouse, move the marker to indicate whether you considered your partner's scores to be very similar, very different, or somewhat different than your own scores." A one-way ANOVA with this reported discrepancy after Trial 1 as the dependent variable indicated a significant difference across discrepancy conditions, $F(2,51) = 59.31$, $p < .001$. Post-hoc tests (Tukey, Sheffe, and Bonferroni) indicated significant differences between low and moderate discrepancy, moderate and high discrepancy, and low and high discrepancy. Likewise, a one-way ANOVA with reported discrepancy on Trial 2 as the dependent variable also indicated a significant difference across discrepancy conditions, $F(2,50) = 76.48$, $p < .001$. One student in the high discrepancy condition did not answer this question. Post-hoc tests (Tukey, Sheffe, and Bonferroni) indicated significant differences between low and moderate discrepancy, moderate and high discrepancy, and low and high discrepancy.

As a second manipulation check, participants were instructed as follows: "Using the scroll bar below, please indicate how you felt about your scores on the first [second] individual trial. Using your mouse, move the marker to indicate whether you considered your scores to be

very low, very high, or somewhere in between", and, "Using the scroll bar below, please indicate how you felt about your partner's scores on the first [second] individual trial. Using your mouse, move the marker to indicate whether you considered your partner's scores to be very low, very high, or somewhere in between." The participants' rating of their own ability was subtracted from their rating of their partner's ability to create a difference measure. A one-way ANOVA with this calculated discrepancy on Trial 1 as the dependent variable indicated a significant difference across discrepancy conditions, $F(2,51) = 21.96, p < .001$. Post-hoc tests (Tukey, Sheffe, and Bonferroni) indicated significant differences between moderate and high discrepancy, and low and high discrepancy, but not between low and moderate discrepancy. A one-way ANOVA with calculated discrepancy on Trial 2 as the dependent variable also indicated a significant difference across discrepancy conditions, $F(2,50) = 22.11, p < .001$. One student in the high discrepancy condition did not answer these questions. Post-hoc tests (Tukey, Sheffe, and Bonferroni) indicated significant differences between moderate and high discrepancy, and low and high discrepancy, but not between low and moderate discrepancy.

As a third and final discrepancy check, participants were asked in a multiple-choice question to indicate whether the difference between their scores and their partners scores was very low, low, moderate, high, or very high. Responses were as expected by condition. After Trial 1, 87.5% of those in the low discrepancy condition reported the difference as "very low", 88.9% of those in the moderate discrepancy condition reported the difference as "moderate", and 75.0% of those in the high condition reported the difference as "high". After Trial 2, 87.5% of those in the low discrepancy condition reported the difference as "very low", 94.4% of those in the moderate discrepancy condition reported the difference as "moderate", and 89.5% of those in the high condition reported the difference as "high".

It appears that participants in the conjunctive conditions viewed the differences between their scores and their partner's scores as low, moderate, and high, as intended.

Conjunctive Task Check. Participants responded to the manipulation checks for task before the first team trial, after the first team trial (i.e., before the second team trial) and after the second team trial. Responses before the first team trial were as expected: 100% of the participants correctly answered that the team trial stops when one person has stopped; 92.3% of the participants correctly answered that the team score is equal to the score of the first person to stop; and, 96.2% of the participants correctly answered that when one person stops, the other person can not continue.

Responses after the first team trial (i.e., before the second team trial) were also as expected: 100% of the participants correctly answered that the team trial stops when one person has stopped; 96.2% of the participants correctly answered that the team score is equal to the score of the first person to stop; and, 98.1% of the participants correctly answered that when one person stops, the other person can not continue. Finally, responses after the second team trial were also as expected: 100% of the participants correctly answered that the team trial stops when one person has stopped; 96.2% of the participants correctly answered that the team score is equal to the score of the first person to stop; and, 96.2% of the participants correctly answered that when one person stops, the other person can not continue.

Thus, it appears that participants understood the conjunctive nature of the team task.

Discrepancy Check for Additive Conditions. 48 students were in the additive conditions: 16 in the low discrepancy condition, 16 in the moderate discrepancy condition, and 16 in the high discrepancy condition. The discrepancy check measures were the same as those used in the conjunctive conditions.

A one-way ANOVA with reported discrepancy on Trial 1 as the dependent variable indicated a significant difference across discrepancy conditions, $F(2,45) = 110.69$, $p < .001$. Post-hoc tests (Tukey, Sheffe, and Bonferroni) indicated significant differences between low and moderate discrepancy, moderate and high discrepancy, and low and high discrepancy. Likewise, a one-way ANOVA with reported discrepancy on Trial 2 as the dependent variable also indicated a significant difference across discrepancy conditions, $F(2,45) = 42.37$, $p < .001$. Post-hoc tests (Tukey, Sheffe, and Bonferroni) indicated significant differences between low and moderate discrepancy, moderate and high discrepancy, and low and high discrepancy.

A one-way ANOVA with calculated discrepancy on Trial 1 as the dependent variable indicated a significant difference across discrepancy conditions, $F(2,45) = 28.17$, $p < .001$. Post-hoc tests (Tukey, Sheffe, and Bonferroni) indicated significant differences between low and moderate discrepancy, moderate and high discrepancy, and low and high discrepancy. A one-way ANOVA with calculated discrepancy on Trial 2 as the dependent variable also indicated a significant difference across discrepancy conditions, $F(2,45) = 24.27$, $p < .001$. Post-hoc tests (Tukey, Sheffe, and Bonferroni) indicated significant differences between moderate and high discrepancy, and low and high discrepancy, but not between low and moderate discrepancy.

Responses to the multiple-choice question were also as expected by condition. After Trial 1, 75.0% of those in the low discrepancy condition reported the difference as "very low", 75.0% of those in the moderate discrepancy condition reported the difference as "moderate", and 56.3% of those in the high condition reported the difference as "high". After Trial 2, 81.0% of those in the low discrepancy condition reported the difference as "very low", 81.0% of those in the moderate discrepancy condition reported the difference as "moderate", and 68.8% of those in the high condition reported the difference as "high".

Thus, participants appeared to view the difference between their performance and their partner's performance as low, moderate, and high, as intended.

Additive Task Check. Participants responded to the manipulation check for task type before the first team trial, after the first team trial (i.e., before the second team trial) and after the second team trial. Responses before the first team trial were as expected: 97.9% of the participants correctly answered that the team trial continues when one person has stopped; 100% of the participants correctly answered that the team score is equal to the average of the partners' scores on the team trial; and, 97.9% of the participants correctly answered that when one person stops, the other person can continue.

Responses after the first team trial (i.e., before the second team trial) were also as expected: 100% of the participants correctly answered that the team trial continues when one person has stopped; 100% of the participants correctly answered that the team score is equal to the average of the partners' scores on the team trial; and, 97.9% of the participants correctly answered that when one person stops, the other person can continue. Finally, responses after the second team trial were also as expected: 100% of the participants correctly answered that the team trial continues when one person has stopped; 100% of the participants correctly answered that the team score is equal to the average of the partners' scores on the team trial; and, 89.6% of the participants correctly answered that when one person stops, the other person can continue.

Thus, participants appeared to understand the additive nature of the team task.

Discussion

The purpose of this pilot study was to assure that participants in the low, moderate, and high discrepancy conditions viewed the differences between their scores and their partner's scores as low, moderate, or high. Indeed, this was the case. A second purpose of this study was

to assure that participants understood the conjunctive or additive nature of the team task. Again, this appeared to be the case. This assurance that participants viewed the manipulations as intended allowed them to be used in the main data collection with greater confidence.

Appendix C

Questionnaire Used in Main Studies

The conditions under which the questions were administered are designated as follows: T = Team Conditions; C = Coactive Conditions; I = Control Condition; TC = Team and Coactive Conditions; IC = Control and Coactive Conditions; A = All Conditions. Response format is noted in brackets.

Performance Improvement Expectancy Measure administered after Discrepancy Feedback (after Trial 1 and after Trial 2) in Team and Coactive Conditions

TC1 and TC4. You have achieved a score in the 17th place. Assuming other participants' scores remain the same, what is the chance that you will score in that place again? Please indicate your answers as a percentage, ranging from 0% chance to 100% chance. ___%
[Fill in the blank]

TC2 and TC5. Assuming other participants' scores remain the same, what place do you feel you have a 100% chance of scoring in? ____ place
[Fill in the blank]

TC3 and TC6. Considering your answer to question 2, please indicate the chance you have of scoring in each of the following places (again, assuming other participants' scores remain the same). ____ % for 1-34 places
[Fill in the blank]

Performance Improvement Expectancy Measure administered after Trial 1 and after Trial 2 in Control Condition

I1 and I4. What is the chance that you score will remain the same on the next trial? Please indicate your answer as a percentage, ranging from 0% chance to 100% chance. ___%
[Fill in the blank]

I2 and I5. What is the chance that you will improve your score on the next trial? Please indicate your answers as a percentage, ranging from 0% chance to 100% chance. ___%
[Fill in the blank]

I3 and I6. Considering your answer to question 2, please indicate the chance you have of increasing or decreasing your score by the following amounts. 100% increase (double your score), 90% increase, 80% increase, 70% increase, 60% increase, 50% increase, 40% increase, 30% increase, 20% increase 10% increase, stay the same, 10% decrease, 20% decrease, 30%

decrease, 40% decrease, 50% decrease, 60% decrease, 70% decrease, 80% decrease, 90% decrease
[Fill in the blank]

Final Questionnaire Administered after Trial 4

Indispensability measure for Team Conditions

T7. To what extent did the team score depend on your performance?
[Scroll bar with response anchors "The team score did not depend on my performance at all" to "The team score was completely dependent on my performance"]

T8. To what extent did the team score depend on your partner's performance?
[Scroll bar with response anchors "The team score did not depend on my partner's performance at all" to "The team score was completely dependent on my partner's performance"]

Indispensability Measure for Coactive Conditions

C7. To what extent did your score depend on your performance?
[Scroll bar with response anchors "My score did not depend on my performance at all" to "My score was completely dependent on my performance"]

C8. To what extent did the score of the person in the next room depend on your performance?
[Scroll bar with response anchors "His or her score did not depend on my performance at all" to "His or her score was completely dependent on my performance"]

Indispensability Measure for Control Condition

I7. To what extent did your score depend on your performance?
[Scroll bar with response anchors "My score did not depend on my performance at all" to "My score was completely dependent on my performance"]

Discrepancy Manipulation Check for Team Conditions

T9. The difference between your score and your partner's score on the first individual trial is
[Multiple choice: Very Low, Low, Moderate, High, Very High]

T10. Using the scroll bar below, please indicate how you felt about the difference between your scores in the first individual trial and your partner's scores on the first individual trial. Using your mouse, move the marker to indicate whether you considered your partner's scores to be very similar, very different, or somewhat different than your own scores.
[Scroll bar with response anchors "My partner's scores were very similar to my scores" to "My partner's scores were very different from my scores"]

T11. Using the scroll bar below, please indicate how you felt about your scores on the first individual trial. Using your mouse, move the marker to indicate whether you considered your scores to be very low, very high, or somewhere in between.
[Scroll bar with response anchors "My scores were very low" to "My scores were very high"]

T12. Using the scroll bar below, please indicate how you felt about your partner's scores on the first individual trial. Using your mouse, move the marker to indicate whether you considered your partner's scores to be very low, very high, or somewhere in between.
[Scroll bar with response anchors "My partner's were very low" to "My partner's scores were very high"]

T13. The difference between your score and your partner's score on the second individual trial is
[Multiple choice: Very Low, Low, Moderate, High, Very High]

T14. Using the scroll bar below, please indicate how you felt about the difference between your scores in the second individual trial and your partner's scores on the second individual trial. Using your mouse, move the marker to indicate whether you considered your partner's scores to be very similar, very different, or somewhat different than your own scores.
[Scroll bar with response anchors "My partner's scores were very similar to my scores" to "My partner's scores were very different from my scores"]

T15. Using the scroll bar below, please indicate how you felt about your scores on the second individual trial. Using your mouse, move the marker to indicate whether you considered your scores to be very low, very high, or somewhere in between.
[Scroll bar with response anchors "My scores were very low" to "My scores were very high"]

T16. Using the scroll bar below, please indicate how you felt about your partner's scores on the second individual trial. Using your mouse, move the marker to indicate whether you considered your partner's scores to be very low, very high, or somewhere in between.
[Scroll bar with response anchors "My partner's scores were very low" to "My partner's scores were very high"]

Discrepancy Manipulation Check for Coactive Conditions

C9. The difference between your score and the score of the person in the next room on the first performance trial was
[Multiple choice: Very Low, Low, Moderate, High, Very High]

C10. Using the scroll bar below, please indicate how you felt about the difference between your scores in the first performance trial and the scores of the person in the next room on the first performance trial. Using your mouse, move the marker to indicate whether you considered his or her scores to be very similar, very different, or somewhat different than your own scores.
[Scroll bar with response anchors "His/her scores were very similar to my scores" to "His/her scores were very different from my scores"]

C11. Using the scroll bar below, please indicate how you felt about your scores on the first performance trial. Using your mouse, move the marker to indicate whether you considered your scores to be very low, very high, or somewhere in between.
[Scroll bar with response anchors "My scores were very low" to "My scores were very high"]

C12. Using the scroll bar below, please indicate how you felt about the scores of the person in the next room on the first performance trial. Using your mouse, move the marker to indicate whether you considered his or her scores to be very low, very high, or somewhere in between.
[Scroll bar with response anchors "His/her scores were very low" to "His/her scores were very high"]

C13. The difference between your score and the score of the person in the next room on the second performance trial was
[Multiple choice: Very Low, Low, Moderate, High, Very High]

C14. Using the scroll bar below, please indicate how you felt about the difference between your scores in the second performance trial and the scores of the person in the next room on the second performance trial. Using your mouse, move the marker to indicate whether you considered his or her scores to be very similar, very different, or somewhat different than your own scores.
[Scroll bar with response anchors "His/her scores were very similar to my scores" to "His/her scores were very different from my scores"]

C15. Using the scroll bar below, please indicate how you felt about your scores on the second performance trial. Using your mouse, move the marker to indicate whether you considered your scores to be very low, very high, or somewhere in between.
[Scroll bar with response anchors "My scores were very low" to "My scores were very high"]

C16. Using the scroll bar below, please indicate how you felt about the scores of the person in the next room on the second performance trial. Using your mouse, move the marker to indicate whether you considered his or her scores to be very low, very high, or somewhere in between.
[Scroll bar with response anchors "His/her scores were very low" to "His/her scores were very high"]

Effort Measure for Team Conditions

T17. Ignoring your actual performance scores, please indicate the amount of effort you put into the task on the first individual trial.
[Scroll bar with response anchors "I did not try hard at all" to "I tried very hard"]

T18. Ignoring your actual performance scores, please indicate the amount of effort you put into the task on the second individual trial.
[Scroll bar with response anchors "I did not try hard at all" to "I tried very hard"]

T19. Ignoring your actual performance scores, please indicate the amount of effort you put into the task on the first team trial.
[Scroll bar with response anchors "I did not try hard at all" to "I tried very hard"]

T20. Ignoring your actual performance scores, please indicate the amount of effort you put into the task on the second team trial.
[Scroll bar with response anchors "I did not try hard at all" to "I tried very hard"]

Effort Measure for Coactive Conditions and Control Condition

IC17. Ignoring your actual performance scores, please indicate the amount of effort you put into the task on the first performance trial.
[Scroll bar with response anchors "I did not try hard at all" to "I tried very hard"]

IC18. Ignoring your actual performance scores, please indicate the amount of effort you put into the task on the second performance trial.
[Scroll bar with response anchors "I did not try hard at all" to "I tried very hard"]

IC19. Ignoring your actual performance scores, please indicate the amount of effort you put into the task on the third performance trial.
[Scroll bar with response anchors "I did not try hard at all" to "I tried very hard"]

IC20. Ignoring your actual performance scores, please indicate the amount of effort you put into the task on the 4th performance trial.
[Scroll bar with response anchors "I did not try hard at all" to "I tried very hard"]

Enjoyment Measure for All Conditions

A21. Ignoring your actual performance scores, please indicate how much you enjoyed the task.
[Scroll bar with response anchors "I did not enjoy the task at all" to "I enjoyed the task very much"]

Difficulty Measure for All Conditions

A22. Ignoring your actual performance scores, please indicate how difficult you thought the task was.
[Scroll bar with response anchors "The task was not difficult at all" to "The task was very difficult"]

Competition Measure for Team Conditions

T23. To what extent were you competing with your partner, trying to do better than your partner?
[Scroll bar with response anchors "I was not trying to do better than my partner" to "I was trying very hard to do better than my partner"]

Competition Measure for Coactive Conditions

C23. To what extent were you competing with the person in the next room, trying to do better than him or her?
[Scroll bar with response anchors "I was not trying to do better than him or her" to "I was trying very hard to do better than him or her"]

Importance Measure for All Conditions

A24. How personally important was it to you to perform well?
[Scroll bar with response anchors "It was not at all important" to "It was very important"]

Team Importance Measure for Team Conditions

T25. How personally important was it to you that your team perform well?
[Scroll bar with response anchors "It was not at all important" to "It was very important"]

T26. How much did you feel a part of your team?
[Scroll bar with response anchors "I did not feel part of my team at all" to "I very much felt part of my team"]

Relative Ability Measure for Team Conditions

T27. Please rate your own ability on the task compared to your partner's ability on the task.
[Scroll bar with response anchors "My ability is much lower" to "My ability is much higher"]

Relative Ability Measure for Coactive Conditions

C27. Please rate your own ability on the task compared to the ability of the person in the next room.
[Scroll bar with response anchors "My ability is much lower" to "My ability is much higher"]

Lottery Value Measure for All Conditions

A28. Fifty dollars is
[Scroll bar with response anchors "a little money." to "a lot of money."]

A29. How valuable is fifty dollars to you?
[Scroll bar with response anchors "Not valuable" to "Very valuable"]

A30. How much do you want to win the fifty dollars?
[Scroll bar with response anchors "Not much" to "Very much"]

Performance Goal Measures for Team Conditions

T31. Did you have a personal performance goal for the first individual trial?

[Multiple choice: "Yes", "No"]

T32. If yes, what was that goal [what score did you want]?
[Fill in the blank]

T33. How close did you come to meeting that goal?
[Scroll bar with response anchors "I did much worse than my goal," "I met my goal," "I did much better than my goal"]

T34. Did you have a personal performance goal for the second individual trial?
[Multiple choice: "Yes", "No"]

T35. If yes, what was that goal [what score did you want]?
[Fill in the blank]
T36. How close did you come to meeting that goal?
[Scroll bar with response anchors "I did much worse than my goal," "I met my goal," "I did much better than my goal"]

T37. Did you have a personal performance goal for the first team trial?
[Multiple choice: "Yes", "No"]

T38. If yes, what was that goal [what score did you want]?
[Fill in the blank]
T39. How close did you come to meeting that goal?
[Scroll bar with response anchors "I did much worse than my goal," "I met my goal," "I did much better than my goal"]

T40. Did you have a personal performance goal for the second team trial?
[Multiple choice: "Yes", "No"]

T41. If yes, what was that goal [what score did you want]?
[Fill in the blank]
T42. How close did you come to meeting that goal?
[Scroll bar with response anchors "I did much worse than my goal," "I met my goal," "I did much better than my goal"]

Performance Goal Measures for Coactive Conditions and Control Condition

IC31. Did you have a personal performance goal for the first performance trial?
[Multiple choice: "Yes", "No"]

IC32. If yes, what was that goal [what score did you want]?
[Fill in the blank]
IC33. How close did you come to meeting that goal?

[Scroll bar with response anchors "I did much worse than my goal," "I met my goal," "I did much better than my goal"]

IC34. Did you have a personal performance goal for the second performance trial?
[Multiple choice: "Yes", "No"]

IC35. If yes, what was that goal [what score did you want]?
[Fill in the blank]
IC36. How close did you come to meeting that goal?
[Scroll bar with response anchors "I did much worse than my goal," "I met my goal," "I did much better than my goal"]

IC37. Did you have a personal performance goal for the third performance trial?
[Multiple choice: "Yes", "No"]

IC38. If yes, what was that goal [what score did you want]?
[Fill in the blank]
IC39. How close did you come to meeting that goal?
[Scroll bar with response anchors "I did much worse than my goal," "I met my goal," "I did much better than my goal"]

IC40. Did you have a personal performance goal for the fourth performance trial?
[Multiple choice: "Yes", "No"]

IC41. If yes, what was that goal [what score did you want]?
[Fill in the blank]
IC42. How close did you come to meeting that goal?
[Scroll bar with response anchors "I did much worse than my goal," "I met my goal," "I did much better than my goal"]

Distraction Measure 1 for Team Conditions

T43. Did you wear your earmuffs? Please be honest! We need to know which participants did not wear the earmuffs because it may influence concentration and performance on the task.

T44. I wore earmuffs on the first individual trial.
[Multiple choice: "Yes", "No"]

T45. I wore earmuffs on the second individual trial.
[Multiple choice: "Yes", "No"]

T46. I wore earmuffs on the first team trial.
[Multiple choice: "Yes", "No"]

T47. I wore earmuffs on the second team trial.
[Multiple choice: "Yes", "No"]

T48. Please describe any distractions you encountered. Which trial were you distracted?
[Open response]

Distraction Measure 1 for Coactive Conditions and Control Condition

IC43. Did you wear your earmuffs? Please be honest! We need to know which participants did not wear the earmuffs because it may influence concentration and performance on the task.

IC44. I wore earmuffs on the first performance trial.
[Multiple choice: "Yes", "No"]

IC45. I wore earmuffs on the second performance trial.
[Multiple choice: "Yes", "No"]

IC46. I wore earmuffs on the third performance trial.
[Multiple choice: "Yes", "No"]

IC47. I wore earmuffs on the fourth performance trial.
[Multiple choice: "Yes", "No"]

IC48. Please describe any distractions you encountered. Which trial were you distracted?
[Open response]

Distraction Measure 2 for All Conditions

A49. Do you feel your performance was affected by distraction?
[Multiple choice: "Yes", "No"]

A50. If yes, was your performance better or worse because of the distraction?
[Multiple choice: "Better", "Worse"]

A51. If yes, please indicate the extent to which your performance was affected.
[Scroll bar response anchors "My performance was slightly affected," "My performance was greatly affected"]

A52. If yes, please indicate on which trials you feel your performance was affected.
[Multiple choice (more than one could be chosen): "First Trial", "Second Trial," "Third Trial," "Fourth Trial"]

Gaming Experience Measure for All Conditions

A53. How often do you play video games that involve first-person shooting (such as Doom, Quake)?
[Multiple choice: "Never," "A few times a year," "A few times a month," "A few times a week," "Every day"]

Task Manipulation Check for Team Conditions

T54. Who was the first to stop during the first team trial?
[Multiple choice, "I stopped first," "My partner stopped first"]

T55. Who was the first to stop during the second team trial?
[Multiple choice, "I stopped first," "My partner stopped first"]

T56. Which of the following statements is true?
[Multiple choice, "During the team trial, when one person stops the other person can NOT continue" "During the team trial, when one person stops the other person CAN continue"]

T57. What happens to the team trial when one person has stopped (has three strikes)?
[Multiple choice, "It continues until the other person stops," "It ends for both people"]

T58. The team score is equal to
[Multiple choice, "the score of the first person to stop during the team trial," "the average of the scores of the first person to stop and the second person to stop during the team trial," "the score of the second person to stop during the team trial"]

Demographic Questions for All Conditions

A59. Please indicate your gender.
[Multiple choice, "Male," "Female"]

A60. Please indicate your age.
[Fill in the blank]

Comments Question for All Conditions

A61. Please list any additional comments you have regarding the task or experiment.
[Open ended]

Table 1

Order of Operations on the Computer for Participants in the Collective Conditions

Sequence	Operation	Description
1	Practice Trial	10 numbers presented
2	Instruction	Told to go back to the middle room
3	Instruction	Reminders regarding learning trial
4	Prompt	Type in "learning Trial"
5	Learning Trial	Max 50 numbers presented
6	Instruction	Reminders regarding first individual trial
7	Prompt	Type in "first individual trial"
8	Prompt	Enter an identification code
9	Prompt	Told that code taken, enter another code*
10	Individual Trial	Continued until 3 strikes accumulated
11	Instruction	Reminders regarding second individual trial
12	Instruction	Feedback introduced
13	Feedback	List of ostensible scores, including "partner's score"
14	Questionnaire	Expectancy measure
15	Prompt	Type in "second individual trial"
16	Individual Trial	Continued until 3 strikes accumulated
17	Instruction	Reminders regarding first team trial
18	Instruction	Feedback introduced

19	Feedback	List of ostensible scores, including "partner's score"
20	Questionnaire	Expectancy measure
21	Prompt	Type in "first team trial"
22	Prompt	Type in team name
23	Team Trial	Continued until 3 strikes accumulated
24	Instruction	Reminders regarding second team trial
25	Scores	Individual scores, team score, lottery tickets
26	Prompt	Type in "second team trial"
27	Prompt	Type in team name
28	Team Trial	Continued until 3 strikes accumulated
29	Scores	Individual scores, team score, lottery tickets
30	Questionnaire	Final Questionnaire
31	Exit	Exit program

*To avoid participants sharing their ID code with prospective participants, and to enhance the believability of the list of scores ostensibly from previous participants, the first ID code entered by the participant was noted as "already taken", and participants had to enter a second ID code. The "already taken" code then later appeared in the list of ostensible scores from previous participants.

Table 2

Order of Operations on the Computer for Participants in the Coactive Conditions

Sequence	Operation	Description
1	Practice Trial	10 numbers presented
2	Instruction	Told to go back to the middle room
3	Instruction	Reminders regarding learning trial
4	Prompt	Type in "learning trial"
5	Learning Trial	Max 50 numbers presented
6	Instruction	Reminders regarding first performance trial
7	Prompt	Type in "first performance trial"
8	Prompt	Enter an identification code
9	Prompt	Told that code taken, enter another code*
10	Performance Trial	Continued until 3 strikes accumulated
11	Instruction	Reminders regarding second performance trial
12	Instruction	Feedback introduced
13	Feedback	List of ostensible scores, including "other's score"
14	Questionnaire	Expectancy measure
15	Prompt	Type in "second performance trial"
16	Performance Trial	Continued until 3 strikes accumulated
17	Instruction	Reminders regarding third performance trial
18	Instruction	Feedback introduced

19	Feedback	List of ostensible scores, including "other's score"
20	Questionnaire	Expectancy measure
21	Prompt	Type in "third performance trial"
22	Performance Trial	Continued until 3 strikes accumulated
23	Instruction	Reminders regarding fourth performance trial
24	Score	Participant's score and lottery tickets
25	Prompt	Type in "fourth performance trial"
26	Performance Trial	Continued until 3 strikes accumulated
27	Score	Participant's score and lottery tickets
28	Questionnaire	Final Questionnaire
29	Exit	Exit program

*To avoid participants sharing their ID code with prospective participants, and to enhance the believability of the list of scores ostensibly from previous participants, the first ID code entered by the participant was noted as "already taken", and participants had to enter a second ID code. The "already taken" code then later appeared in the list of ostensible scores from previous participants.

Table 3

Order of Operations on the Computer for Participants in the Control Condition

Sequence	Operation	Description
1	Practice Trial	10 numbers presented
2	Instruction	Told to go back to the middle room
3	Instruction	Reminders regarding learning trial
4	Prompt	Type in "learning trial"
5	Learning Trial	Max 50 numbers presented
6	Instruction	Reminders regarding first performance trial
7	Prompt	Type in "first performance trial"
8	Performance Trial	Continued until 3 strikes accumulated
9	Instruction	Reminders regarding second performance trial
10	Questionnaire	Expectancy measure
11	Prompt	Type in "second performance trial"
12	Performance Trial	Continued until 3 strikes accumulated
13	Instruction	Reminders regarding third performance trial
14	Questionnaire	Expectancy measure
15	Prompt	Type in "third performance trial"
16	Performance Trial	Continued until 3 strikes accumulated
17	Instruction	Reminders regarding fourth performance trial
18	Score	Participant's score and lottery tickets

19	Prompt	Type in "fourth performance trial"
20	Performance Trial	Continued until 3 strikes accumulated
21	Score	Participant's score and lottery tickets
22	Questionnaire	Final Questionnaire
23	Exit	Exit program

Table 4.

Number of Participants by Condition in Study 1.

	Discrepancy Level			
	Low	Moderate	High	Total
Conjunctive	15 males	12 males	13 males	40 males
	18 females	16 females	13 females	47 females
	33 total	28 total	26 total	87 total
Coactive	16 males	16 males	15 males	47 males
	16 females	15 females	16 females	47 females
	32 total	31 total	31 total	94 total
Total	31 males	28 males	28 males	87 males
	34 females	31 females	29 females	94 females
	65 total	59 total	57 total	181 total

Table 5.

Percentage of Responses to Discrepancy Manipulation Check for Study 1.

	Response				
	Very Low	Low	Moderate	High	Very High
Trial 1					
Conjunctive					
Low	72.7	6.1	15.2	3.0	0
Moderate	3.6	3.6	82.1	3.6	0
High	0	3.8	7.7	80.8	7.7
Coactive					
Low	71.9	15.6	9.4	0	3.1
Moderate	0	3.2	90.3	0	6.5
High	0	0	11.5	61.5	26.9
Trial 2					
Conjunctive					
Low	60.6	9.1	12.1	3.0	15.2
Moderate	0	14.3	57.1	0	28.6
High	0	0	11.5	61.5	26.9
Coactive					
Low	62.5	12.5	9.4	0	15.6

Moderate	0	0	80.6	3.2	16.1
High	0	0	3.2	74.2	22.6

Note: The question for the conjunctive conditions was: "The difference between your score and the score of your partner on the first (second) performance trial was....". The question for the coactive conditions was: "The difference between your score and the score of the person in the next room on the first (second) performance trial was....".

In the conjunctive/low discrepancy condition 3.0% were missing data, and in the conjunctive/moderate condition 7.1% were missing data.

Table 6.

Raw Performance Means and Standard Deviations from Study 1 by Trial and Condition.

		Discrepancy Level			
		Low	Moderate	High	Total
Trial 1					
	Conjunctive	143.03	176.12	157.19	157.91
		(64.97)	(84.33)	(79.23)	(76.30)
	Coactive	178.66	181.74	168.64	176.37
		(84.01)	(99.10)	(76.28)	(86.19)
	Total	160.57	179.07	163.42	167.50
		(76.50)	(91.64)	(77.15)	(81.88)
Trial 2					
	Conjunctive	161.94	214.00	169.42	180.93
		(63.35)	(118.11)	(77.70)	(90.35)
	Coactive	214.16	184.42	197.13	198.73
		(140.13)	(106.39)	(131.26)	(126.12)
	Total	187.65	198.46	184.49	190.18
		(110.50)	(112.12)	(110.08)	(110.44)
Trial 4					
	Conjunctive	174.21	270.54	219.04	218.61
		(97.98)	(178.33)	(166.95)	(152.62)
	Coactive	206.66	239.97	251.26	232.35

	(153.34)	(162.93)	(145.88)	(153.72)
Total	190.18	254.47	236.56	225.75
	(128.28)	(169.62)	(155.26)	(152.92)

Note: Standard deviations are in parentheses.

Table 7.

Transformed Performance Means and Standard Deviations from Study 1 by Trial and Condition.

		Discrepancy Level			
		Low	Moderate	High	Total
Trial 1					
	Conjunctive	2.12	2.20	2.15	2.15
		(.19)	(.20)	(.20)	(.20)
	Coactive	2.21	2.20	2.19	2.20
		(.19)	(.22)	(.18)	(.20)
	Total	2.16	2.20	2.17	2.18
		(.19)	(.21)	(.19)	(.20)
Trial 2					
	Conjunctive	2.18	2.27	2.13	2.20
		(.16)	(.24)	(.40)	(.28)
	Coactive	2.26	2.19	2.18	2.21
		(.27)	(.31)	(.36)	(.31)
	Total	2.22	2.23	2.16	2.20
		(.22)	(.28)	(.38)	(.30)
Trial 4					
	Conjunctive	2.02	2.20	2.09	2.10
		(.25)	(.23)	(.26)	(.26)
	Coactive	2.06	2.15	2.16	2.12

	(.31)	(.21)	(.26)	(.27)
Total	2.04	2.17	2.13	2.11
	(.28)	(.22)	(.26)	(.26)

Note: Standard deviations are in parentheses. Trial 1 and Trial 2 scores are log transformed.

Trial 4 scores are log transformed and corrected for practice, fatigue and lottery effects.

Table 8.

Correlations in Conjunctive Conditions from Study 1.

	Disc.	Ability	Indisp.	Goal	Expect.
Disc.					
Ability	-.232*				
Indisp	.077	-.222*			
Goal	.002	-.238	.091		
Expect.	.001	.122	-.163	-.080	
Perf.	.086	-.048	-.105	.786**	-.030

*significant at .05 level

**significant at .01 level

Note: Pairwise deletion.

Table 9.

Correlations in Coactive Conditions from Study 1.

	Disc.	Ability	Indisp.	Goal	Expect.
Disc.					
Ability	-.334**				
Indisp	.058	-.032			
Goal	.170	-.262	-.099		
Expect.	.107	.011	-.099	-.126	
Perf.	.116	.042	.022	.490**	-.021

*significant at .05 level

**significant at .01 level

Note: Pairwise deletion.

Table 10.

Number of Participants by Condition in Study 2.

	Discrepancy Level			
	Low	Moderate	High	Total
Additive	9 males	10 males	15 males	34 males
	12 females	15 females	16 females	43 females
	21 total	25 total	31 total	77 total
Coactive	15 males	18 males	16 males	49 males
	14 females	16 females	16 females	46 females
	29 total	34 total	32 total	95 total
Total	24 males	28 males	31 males	83 males
	26 females	31 females	32 females	89 females
	50 total	59 total	63 total	172 total

Table 11.

Percentage of Responses to Discrepancy Manipulation Check for Study 2.

	Response				
	Very Low	Low	Moderate	High	Very High
Trial 1					
Additive					
Low	61.9	23.8	14.3	0	0
Moderate	4.0	12.0	72.0	12.0	0
High	0	0	16.1	67.7	16.1
Coactive					
Low	72.4	17.2	6.9	3.4	0
Moderate	0	2.9	94.1	2.9	0
High	0	3.1	9.4	81.3	6.3
Trial 2					
Additive					
Low	57.1	14.3	9.5	4.8	14.3
Moderate	4.0	4.0	76.0	4.0	12.0
High	0	3.2	9.7	45.2	41.9
Coactive					
Low	58.6	13.8	6.9	3.4	17.2

Moderate	0	2.9	70.6	5.9	20.6
High	17.4	6.4	30.8	23.8	21.5

Note: The question for the additive conditions was: "The difference between your score and the score of your partner on the first (second) performance Trial was....". The question for the coactive conditions was: "The difference between your score and the score of the person in the next room on the first (second) performance Trial was....".

Table 12.

Raw Performance Means and Standard Deviations from Study 2 by Trial and Condition.

		Discrepancy Level			
		Low	Moderate	High	Total
Trial 1					
	Additive	147.71	176.08	205.52	180.19
		(67.20)	(77.58)	(121.64)	(97.41)
	Coactive	166.62	161.03	177.22	168.19
		(71.84)	(85.20)	(95.16)	(84.37)
	Total	158.68	167.41	191.14	173.56
		(69.87)	(81.71)	(109.04)	(90.37)
Trial 2					
	Additive	171.71	190.12	207.84	192.23
		(58.39)	(92.57)	(71.78)	(76.51)
	Coactive	163.55	224.38	235.00	209.39
		(77.17)	(202.75)	(171.60)	(163.92)
	Total	166.98	209.86	221.63	201.71
		(69.36)	(165.00)	(131.93)	(132.08)
Trial 4					
	Additive	250.05	244.80	265.90	254.73
		(112.44)	(140.58)	(195.58)	(157.34)

Coactive	224.97	279.18	289.59	266.14
	(93.86)	(233.39)	(210.15)	(192.57)
Total	235.50	264.41	277.94	261.03
	(101.74)	(198.65)	(201.82)	(177.26)

Note: Standard deviations are in parentheses.

Table 13.

Transformed Performance Means and Standard Deviations from Study 2 by Trial and Condition.

		Discrepancy Level			
		Low	Moderate	High	Total
Trial 1					
	Additive	2.13	2.21	2.23	2.20
		(.19)	(.18)	(.30)	(.24)
	Coactive	2.19	2.15	2.20	2.18
		(.18)	(.23)	(.22)	(.21)
	Total	2.16	2.17	2.21	2.19
		(.18)	(.21)	(.26)	(.22)
Trial 2					
	Additive	2.21	2.24	2.29	2.25
		(.17)	(.20)	(.16)	(.18)
	Coactive	2.15	2.23	2.28	2.22
		(.25)	(.35)	(.29)	(.30)
	Total	2.18	2.23	2.29	2.24
		(.22)	(.29)	(.23)	(.25)
Trial 4					
	Additive	2.20	2.17	2.17	2.18
		(.17)	(.20)	(.27)	(.22)
	Coactive	2.16	2.17	2.20	2.18

	(.15)	(.32)	(.27)	(.26)
Total	2.18	2.17	2.19	2.18
	(.16)	(.27)	(.27)	(.24)

Note: Standard deviations are in parentheses. Trial 1 and Trial 2 scores are log transformed. Trial 4 scores are log transformed and corrected for practice, fatigue, and lottery effects.

Table 14.

Correlations in Additive Conditions from Study 2.

	Disc.	Ability	Indisp.	Goal	Expect.
Disc.					
Ability	.266*				
Indisp	.386**	.519**			
Goal	.147	-.101	.038		
Expect.	-.032	-.012	-.049	.208	
Perf.	-.056	.166	.098	.263	-.049

*significant at .05 level

**significant at .01 level

Note: Pairwise deletion.

Table 15.

Correlations in Coactive Conditions from Study 2.

	Disc.	Ability	Indisp.	Goal	Expect.
Disc.					
Ability	.269*				
Indisp	-.140	.207			
Goal	.196	.024	-.543**		
Expect.	.038	.236*	.059	.051	
Perf.	.090	.178	.004	.409**	.096

*significant at .05 level

**significant at .01 level

Note: Pairwise deletion.

Figure 1.

The discrepancy model.

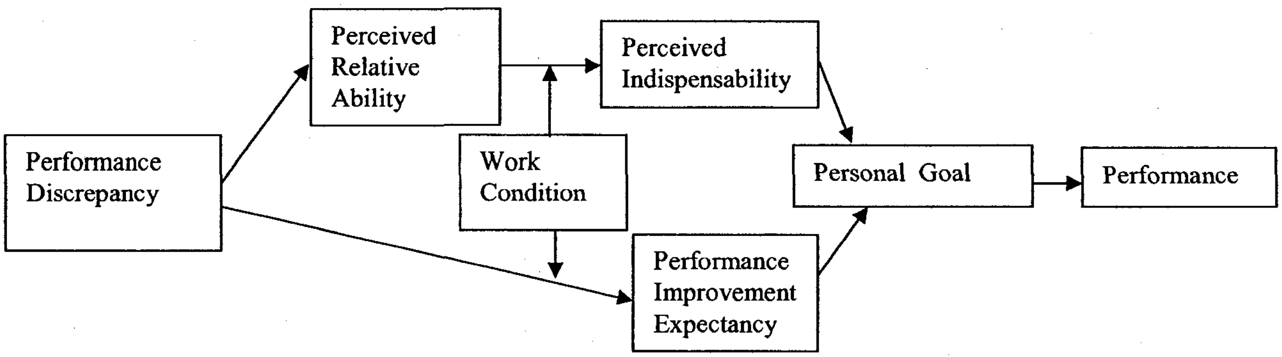


Figure 2.

The discrepancy manipulation for Study 1. Each score in the list is the product of the participant's score and a discrepancy amount. Boldface indicates columns viewed by participants. Italics specifies condition. The partner's id code was always "backdoor". The partner's id and score, and the participant's id and score, were highlighted in red. In this example, the participant's score was 200.

Discrepancy	**ID**	**Score**	
1.87	**0402**	**374**	
1.84	**comp08** ***(backdoor)***	**368**	**Your partner** ***(high discrepancy)***
1.80	**amc**	**360**	
1.68	**seinfeld**	**336**	
1.54	**hot100**	**308**	
1.42	**cubby**	**284**	
1.30	**brag043**	**260**	
1.26	**fish**	**252**	
1.23	**stache02**	**246**	
1.20	**teach05**	**240**	
1.17	**alarm** ***(backdoor)***	**234**	**Your partner** ***(moderate discrepancy)***
1.14	**eyes2**	**228**	
1.11	**music99**	**222**	
1.08	**Barbie**	**216**	
1.05	**shy740** ***(backdoor)***	**210**	**Your partner** ***(low discrepancy)***
1.01	**sunshine**	**202**	
1.00	**(participant's id code)**	**200**	**You**
.98	**jerk040**	**196**	
.97	**sweetie**	**194**	
.90	**051**	**180**	
.84	**repub**	**168**	
.80	**("already taken" code)**	**160**	
.76	**1992**	**152**	
.72	**arman4**	**144**	
.68	**printer**	**136**	
.64	**cheesy**	**128**	
.60	**canoe3**	**120**	
.55	**bb8**	**110**	
.50	**cool99**	**100**	
.45	**party**	**90**	
.40	**nice09**	**80**	
.35	**hvac**	**70**	
.30	**spike**	**60**	
.25	**2001**	**50**	

Your place in the list is	**17**	
Your partner's place in the list is	**15**	*high discrepancy*
	11	*moderate discrepancy*
	2	*low discrepancy*
The difference between your place and your partner's place is	**very low**	*low discrepancy*
	moderate	*moderate discrepancy*
	very high	*high discrepancy*

Figure 3.

Trial 2 and Trial 4 adjusted means by level of discrepancy in Study 1.

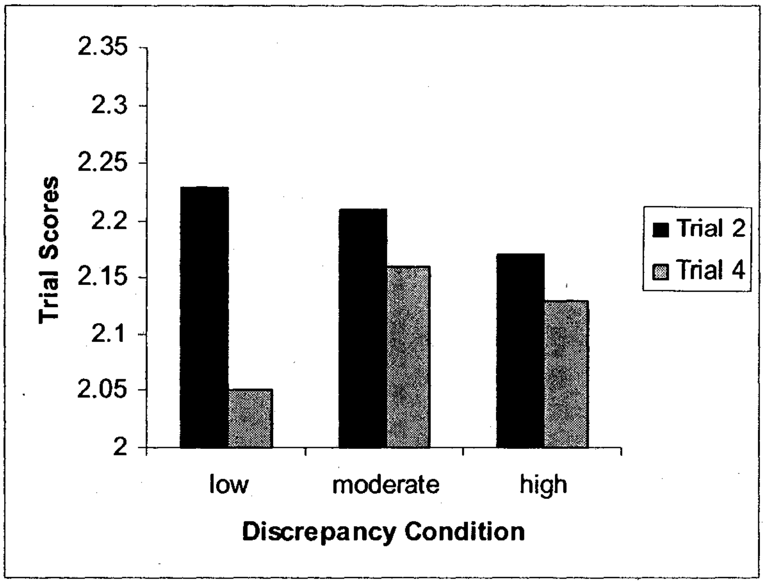
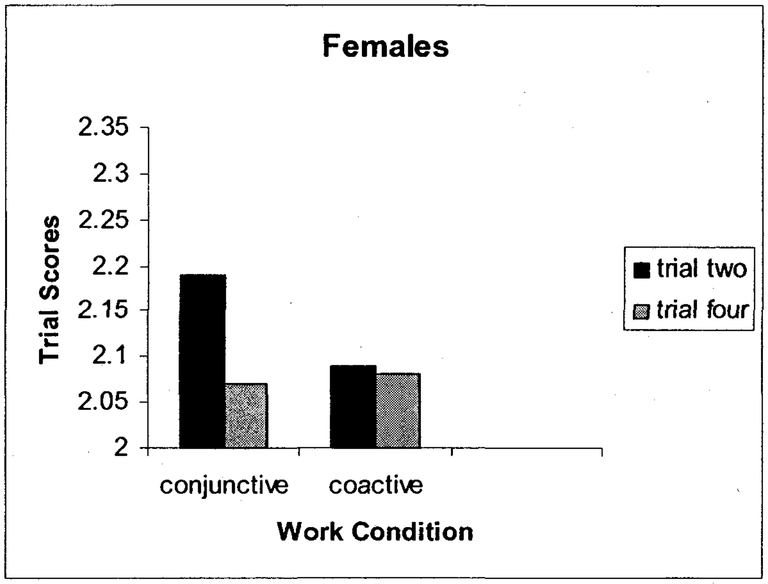


Figure 4.

Trial 2 and Trial 4 unadjusted means by gender and work condition in Study 1.



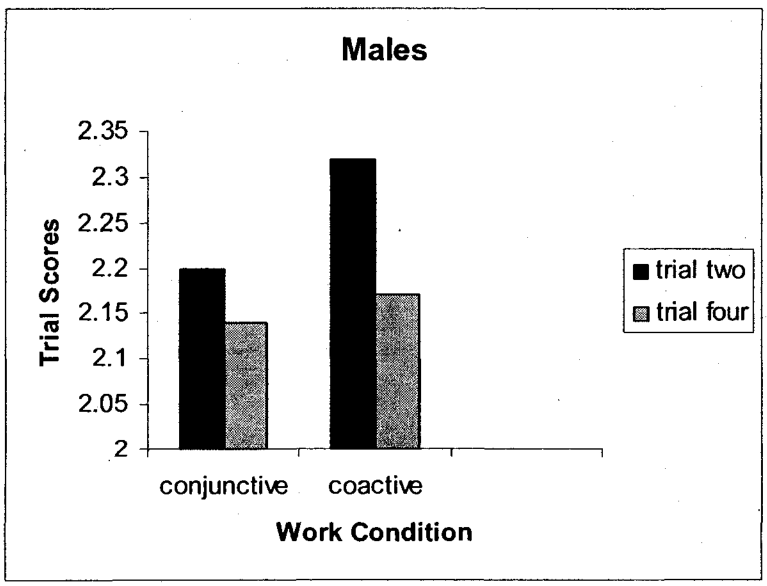


Figure 5.

The experimental model.

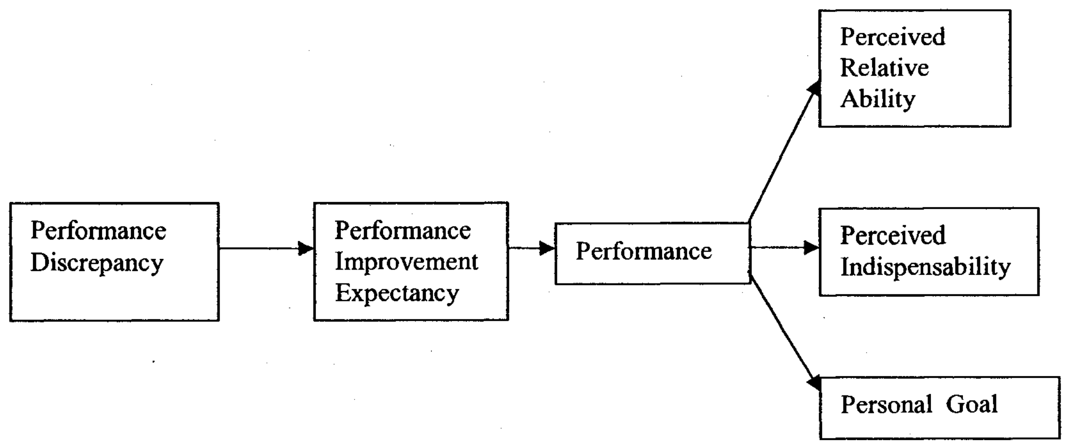


Figure 6.

*Significant path coefficients for conjunctive and coactive groups in the theoretical model. Coefficients for the conjunctive group are in bold. *significant at .05 level, **significant at .01 level.*

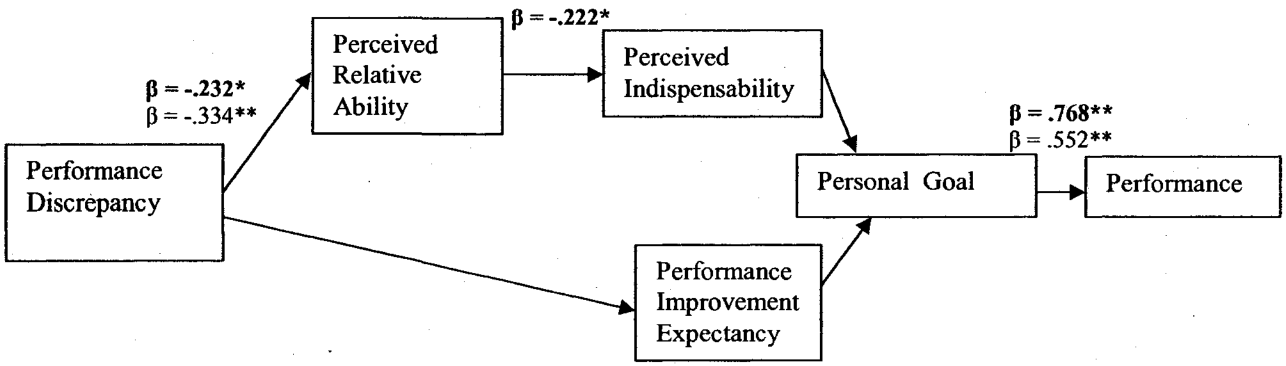


Figure 7.

The discrepancy manipulation for Study 2. Each score in the list is the product of the participant's score and a discrepancy amount. Boldface indicates columns viewed by participants. Italics specifies condition. The partner's id code was always "backdoor". The partner's id and score, and the participant's id and score, were highlighted in red. In this example, the participant's score was 200.

Discrepancy	**ID**	**Score**	
1.87	**0402**	**374**	
1.84	**comp08**	**368**	
1.80	**amc**	**360**	
1.68	**seinfeld**	**336**	
1.54	**hot100**	**308**	
1.42	**cubby**	**284**	
1.30	**brag043**	**260**	
1.26	**fish**	**252**	
1.23	**stache02**	**246**	
1.20	**teach05**	**240**	
1.17	**alarm**	**234**	
1.14	**eyes2**	**228**	
1.11	**music99**	**222**	
1.08	**Barbie**	**216**	
1.05	**shy740**	**210**	
1.01	**sunshine**	**202**	
1.00	**(participant's id code)**	**200**	**You**
.98	**jerk040**	**196**	
.97	**sweetie** ***(backdoor)***	**194**	**Your partner** ***(low discrepancy)***
.90	**051**	**180**	
.84	**repub**	**168**	
.80	**("already taken" code)**	**160**	
.76	**1992**	**152**	
.72	**arman4** ***(backdoor)***	**144**	**Your partner** ***(moderate discrepancy)***
.68	**printer**	**136**	
.64	**cheesy**	**128**	
.60	**canoe3**	**120**	
.55	**bb8**	**110**	
.50	**cool99**	**100**	
.45	**party**	**90**	
.40	**nice09**	**80**	
.35	**hvac** ***(backdoor)***	**70**	**Your partner** ***(high discrepancy)***
.30	**spike**	**60**	
.25	**2001**	**50**	

Your place in the list is	**17**	
Your partner's place in the list is	**32**	*high discrepancy*
	24	*moderate discrepancy*
	19	*low discrepancy*
The difference between your place and your partner's place is	**very low**	*low discrepancy*
	moderate	*moderate discrepancy*
	very high	*high discrepancy*

Figure 8.

Trial 4 error rate adjusted means by gender and discrepancy level in Study 2.

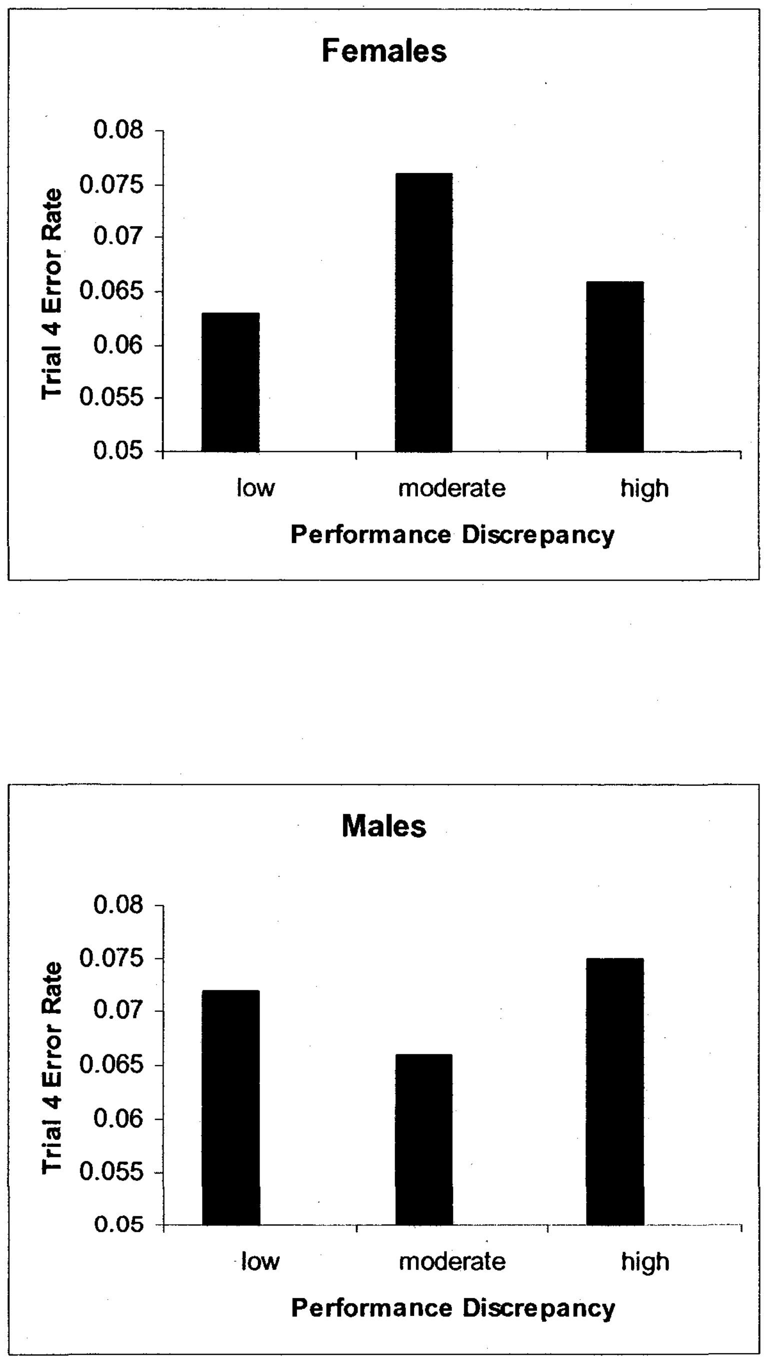


Figure 9.

Perceived indispensability means by discrepancy level from the additive condition in Study 2.

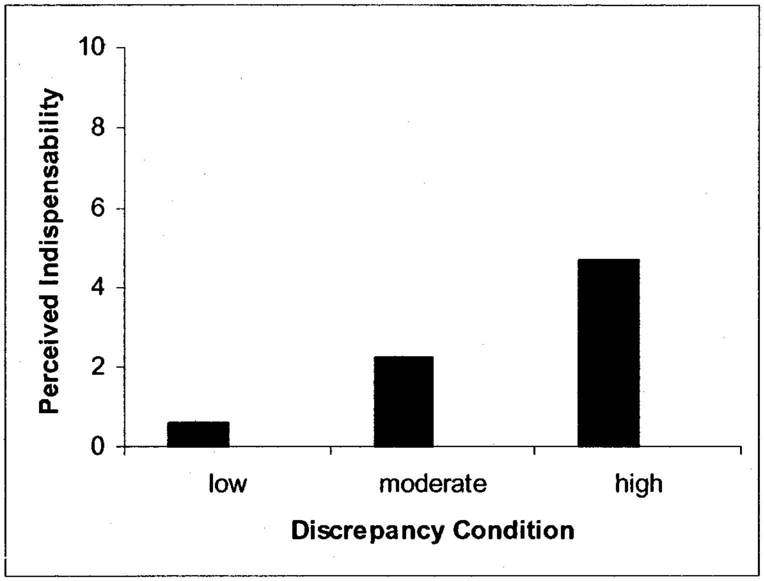
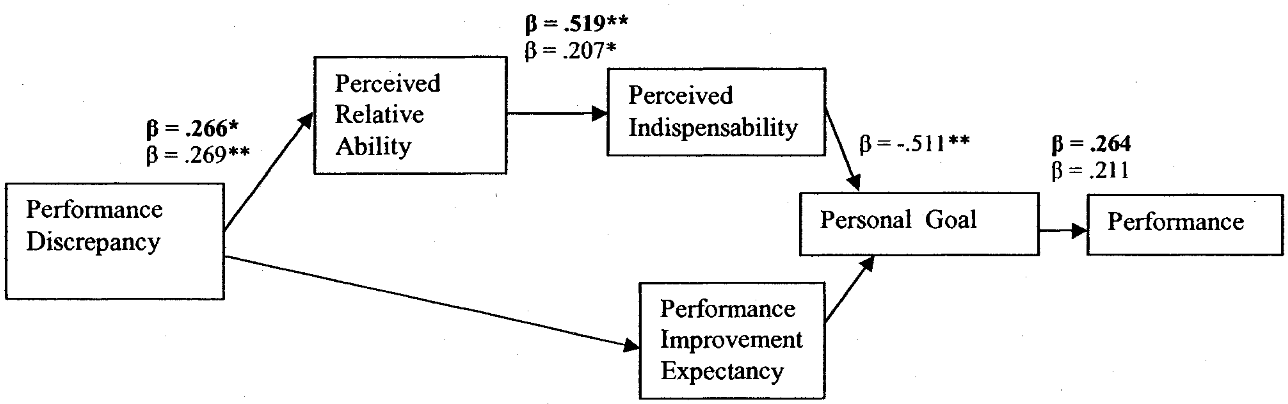


Figure 10.

*Significant path coefficients for additive and coactive groups in the theoretical model. Coefficients for the additive group are in bold. *significant at .05 level, **significant at .01 level.*



UNIVERSITY OF ILLINOIS
AT CHICAGO

Office for the Protection of Research Subjects (OPRS)
Office of the Vice Chancellor for Research (MC 672)
203 Administrative Office Building
1737 West Polk Street
Chicago, Illinois 60612-7227

Approval Notice
Continuing Review

June 13, 2008

Victoria Harmon, MA
Psychology
1009 BSB
1007 W Harrison, M/C 285
Chicago, IL 60612
Phone: (312) 996-3036

RE: Protocol # 2006-0403
"Performance Gain and Motivation Measurement"

Dear Ms. Harmon:

Your Continuing Review Application was reviewed and approved by the Expedited review process on June 10, 2008. You may continue your research.

Please note the following information about your approved research protocol:

The IRB previously approved the enrollment of 600 subjects; however 718 were enrolled. Although no corrective action is required at this time, please remember the importance of not exceeding the IRB-approved subject enrollment number in the future.

Protocol Approval Period: June 19, 2008 - June 18, 2009
Approved Subject Enrollment #: 718 previously enrolled subjects
Performance Sites: UIC
Sponsor: None
Research Protocol(s):

a) Performance Gain and Motivation Measurement

Recruitment Material(s):

a) The remaining research-related activities are limited to data analysis only.

Informed Consent(s):

a) The remaining research-related activities are limited to data analysis only.

Your research meets the criteria for expedited review as defined in 45 CFR 46.110(b)(1) under the following specific category:

(7) Research on individual or group characteristics or behavior (including but not limited to research on perception, cognition, motivation, identity, language, communication, cultural beliefs or practices and social behavior) or research employing survey, interview, oral history, focus group, program evaluation, human factors evaluation, or quality assurance methodologies.

Please note the Review History of this submission:

Receipt Date	Submission Type	Review Process	Review Date	Review Action
06/04/2008	Continuing Review	Expedited	06/10/2008	Approved

Please remember to:

→ Use your **research protocol number** (2006-0403) on any documents or correspondence with the IRB concerning your research protocol.

→ Review and comply with all requirements on the enclosure, "**UIC Investigator Responsibilities, Protection of Human Research Subjects**"

Please note that the UIC IRB has the prerogative and authority to ask further questions, seek additional information, require further modifications, or monitor the conduct of your research and the consent process.

Please be aware that if the scope of work in the grant/project changes, the protocol must be amended and approved by the UIC IRB before the initiation of the change.

We wish you the best as you conduct your research. If you have any questions or need further help, please contact OPRS at (312) 996-1711 or me at (312) 355-2908. Please send any correspondence about this protocol to OPRS at 203 AOB, M/C 672.

Sincerely,

Charles W. Hoehne
Assistant Director, IRB # 2
Office for the Protection of Research Subjects

Enclosure(s): **UIC Investigator Responsibilities, Protection of Human Research Subjects**

cc: Gary E. Raney, Psychology, M/C 285
James R. Larson, Jr., Psychology, M/C 285

UNIVERSITY OF ILLINOIS
AT CHICAGO

Office for the Protection of Research Subjects (OPRS)
Office of the Vice Chancellor for Research (MC 672)
203 Administrative Office Building
1737 West Polk Street
Chicago, Illinois 60612-7227

Approval Notice
Continuing Review

September 9, 2008

Victoria Harmon, MA
Psychology
1009 BSB
1007 W Harrison, M/C 285
Chicago, IL 60612
Phone: (312) 996-3036

RE: Protocol # 2006-0635
"Performance Motivation on a Vigilance Task"

Dear Ms. Harmon:

Your Continuing Review was reviewed and approved by Members of IRB #2 by the Expedited review process on September 4, 2008. You may now continue your research.

Please note the following information about your approved research protocol:

Protocol Approval Period: September 7, 2008 - September 6, 2009
Approved Subject Enrollment #: 100 (79 subjects enrolled; Enrollment closed)
Additional Determinations for Research Involving Minors: The Board determined that this research satisfies 45CFR46.404, research not involving greater than minimal risk
Performance Sites: UIC
Sponsor: Departmental
Research Protocol(s):

a) Performance Motivation on a Vigilance Task

Recruitment Material(s): N/A – Research limited to data analysis only.
Informed Consent(s): N/A – Research limited to data analysis only.
Parental Permission(s): N/A – Research limited to data analysis only.

Your research continues to meet the criteria for expedited review as defined in 45 CFR 46.110(b)(1) under the following specific category:

(7) Research on individual or group characteristics or behavior (including but not limited to research on perception, cognition, motivation, identity, language, communication, cultural beliefs or practices and social behavior) or research employing survey, interview, oral history, focus group, program evaluation, human factors evaluation, or quality assurance methodologies.

Please note the Review History of this submission:

Receipt Date	Submission Type	Review Process	Review Date	Review Action
09/03/2008	Continuing Review	Expedited	09/04/2008	Approved

Please remember to:

→ Use your **research protocol number** (2006-0635) on any documents or correspondence with the IRB concerning your research protocol.

→ Review and comply with all requirements on the enclosure,
"UIC Investigator Responsibilities, Protection of Human Research Subjects"

Please note that the UIC IRB has the prerogative and authority to ask further questions, seek additional information, require further modifications, or monitor the conduct of your research and the consent process.

Please be aware that if the scope of work in the grant/project changes, the protocol must be amended and approved by the UIC IRB before the initiation of the change.

We wish you the best as you conduct your research. If you have any questions or need further help, please contact OPRS at (312) 996-1711 or me at (312) 355-2939. Please send any correspondence about this protocol to OPRS at 203 AOB, M/C 672.

Sincerely,

Jewell Hamilton, MSW
IRB Coordinator, IRB # 2
Office for the Protection of Research Subjects

Enclosure(s):

1. **UIC Investigator Responsibilities, Protection of Human Research Subjects**

cc: Gary E. Raney, Psychology, M/C 285
James R. Larson, Jr., Faculty Sponsor, Psychology, M/C 285

VITA

NAME: Victoria M. Harmon

EDUCATION: B. A., Psychology, Northern Illinois University, DeKalb, Illinois, 1993

M.A., Industrial/Organizational Psychology, Roosevelt University, Chicago, Illinois 1999

TEACHING: Graduate Teaching Assistant, Department of Psychology, University of Illinois at Chicago, Chicago, Illinois, 2001-2008.

Instructor, Statistics for the Behavioral Sciences, Department of Psychology, University of Illinois at Chicago, Chicago, Illinois, 2005.

Instructor, Social Psychology, Department of Psychology, North Central College, Naperville, Illinois, 2005

HONORS: Psychology Departmental Research Grant, University of Illinois at Chicago, 2007

Golden Key National Honor Society, Northern Illinois University,1993

PROFESSIONAL MEMBERSHIP:
Academy of Management
Society for Industrial/Organizational Psychology
American Psychological Society
Midwestern Psychological Association

PRESENTATIONS:

Harmon, V. (2008, August). *Affective/cognitive ambivalence in organizational identification.* Paper presented at the annual meeting of the Academy of Management, Anaheim, Ca.

Harmon, V., & Sargis, E. (2008, July). *Do awareness of information bias and information status in decision-making groups enhance the discussion of unique information?* Poster session presented at the annual meeting of Interdisciplinary Network for Group Research 2008 Meeting, Kansas City, Mo.

VITA (continued)

Harmon, V., & Sargis, E. (2008, May). *Do awareness of information bias and information status in decision-making groups enhance the discussion of unique information?* Poster session presented at the annual meeting of the American Psychological Society, Chicago, Il.

Harmon, V. & Larson, J. (2005, May). *Repeating what we remember: Shared and unshared information in a group discussion.* Poster session presented at the annual meeting of the Midwestern Psychological Association. Chicago, Il.

Harmon, V., Sargis, E., Winquist, J., Franz, T., & Larson, J. (2004, May). *Expectations of group processes influence members' anticipatory coordination.* Poster session presented at the annual meeting of the American Psychological Society, Chicago, Il.

Harmon, V. (2003, May). *Attitude formation in a decision context.* Poster session presented at the annual meeting of the Midwestern Psychological Association. Chicago, Il.

Harmon, V., & Bizot, E. (2000, August). *Combining interests and aptitudes into career recommendations.* Poster session presented at the annual meeting of the American Psychological Association, Washington, D.C.

Harmon, V., Barton, M., & Carson, A. (1999, August). *Further examination of relations between abilities and interest codes.* Poster session presented at the annual meeting of the American Psychological Association, Boston, MA.

PUBLICATIONS:

Larson, J., & Harmon, V. (2007). Recalling shared vs. unshared information mentioned during group discussion: Toward understanding differential repetition rates. *Group Processes and Interpersonal Relations, 10*, 311-322.

O'Donahue, W., Smith*, V., Schewe, P. (1998). The credibility of child sexual abuse allegations: Perpetrator gender and subject occupational status. *Sexual Abuse: A Journal of Research and Treatment, 10*, 17-24. *Maiden name